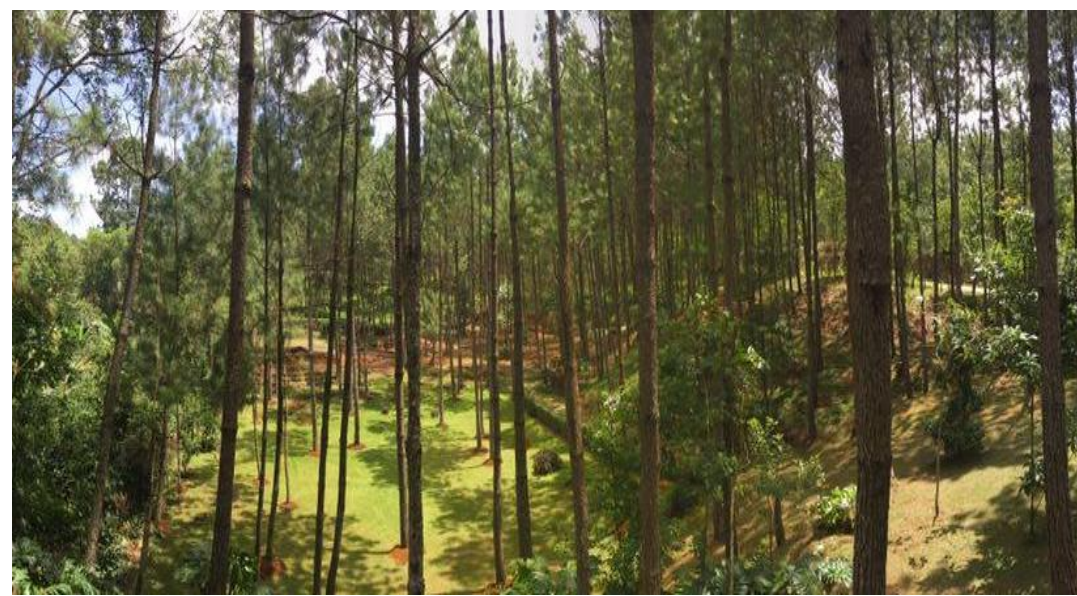
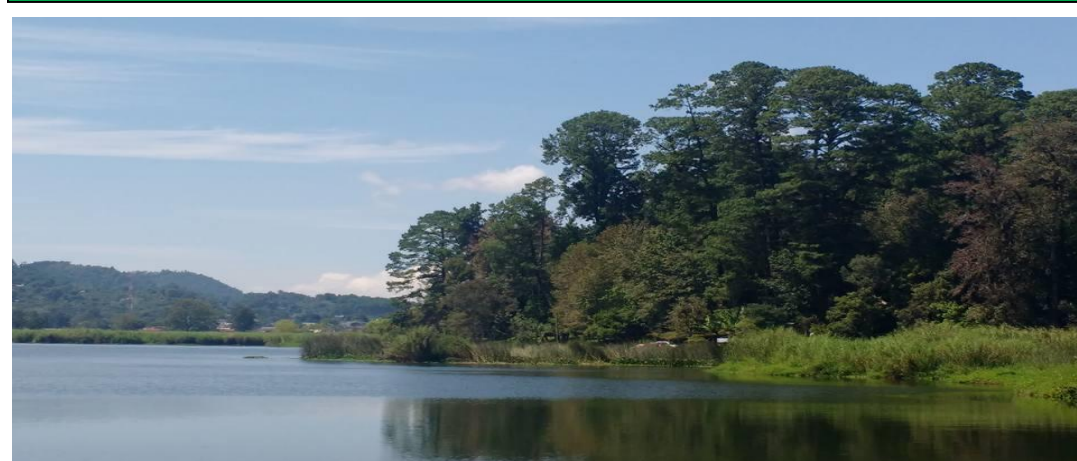
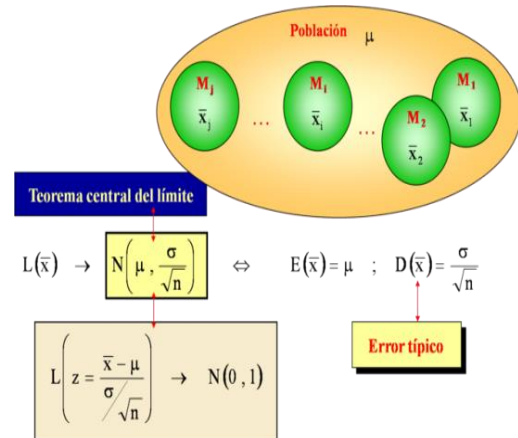




Facultad de Agronomía –Universidad de San Carlos de Guatemala

PRINCIPIOS DE MUESTREO ESTADÍSTICO

Fundamentos y aplicaciones en Agronomía y ciencias afines



Dr. Sc. Ezequiel López
Dr. Byron González

2017

Derechos reservados:

Texto elaborado con fines docentes, se autoriza la reproducción parcial de la información aquí contenida, siempre y cuando se cite esta publicación como fuente.

Subárea de Métodos de Cuantificación e Investigación
Facultad de Agronomía
Universidad de San Carlos de Guatemala
Oficina C-20, Edificio T-8
Ciudad Universitaria zona 12
Ciudad de Guatemala, C.A.

Versión digital y bases de datos disponibles en: cete.fausac.gt

Ciudad de Guatemala, agosto de 2017.



Autores



Ezequiel López es actualmente Profesor Titular VI y Director de la Escuela de Estudios de Postgrado de la Facultad de Agronomía de la Universidad de San Carlos de Guatemala (FAUSAC). Agrónomo egresado de la Escuela Nacional Central de Agricultura –ENCA- (1991), Ingeniero Agrónomo en Sistemas de Producción Agrícola graduado en la FAUSAC (1999), realizó estudios de Maestría en Agronomía (2002) y Doctorado en Ciencias (2014), ambos con Área de concentración en Estadística y Experimentación Agronómica en la Escuela Superior de Agricultura “Luiz de Queiroz” (Piracicaba, SP) de la Universidad de São Paulo, Brasil (USP). Se ha desempeñado como docente de Estadística Aplicada en Agronomía y ciencias afines, en pre y postgrado en la USAC, y en diferentes Universidades Privadas de Guatemala. Además ha sido docente asesor de EPSA, asesor de trabajos de investigación y consultor estadístico en proyectos de investigación. Su área de



Byron González es actualmente Profesor Titular VII, Coordinador de la Sub área de Métodos de Cuantificación e Investigación y Director del Centro de Telemática (CETE) de la Facultad de Agronomía, USAC. Es Doctor en Investigación (2016) por la Universidad Panamericana de Guatemala, realizó estudios de Maestría en Administración de Empresas en la Universidad Rafael Landívar de Guatemala, Ingeniero Agrónomo en Sistemas de Producción Agrícola graduado en la Facultad de Agronomía de la Universidad de San Carlos de Guatemala (1999), Agrónomo egresado de la Escuela Nacional Central de Agricultura (ENCA) 1991. Se ha desempeñado como docente de Informática, Estadística Aplicada en Administración, Economía y Agronomía, en pregrado en la FAUSAC y postgrado en diversas Universidades Privadas de Guatemala. Además como asesor de

Consultas y sugerencias:

e-mails:lopez_ezequiel@usac.edu.gt, ealbautis@gmail.com

byrong@usac.edu.gt, byron.gonzalez@cetegt.org

PRESENTACIÓN

Para un investigador que necesita conocer diferentes aspectos de su interés en una población específica, resulta imposible recurrir a todos los individuos que la conforman para poder cumplir sus objetivos. Por esta razón se hace necesario tomar **muestras** y analizarlas, para que con un esfuerzo y costos razonables se pueda llegar a conclusiones tan válidas como las que habría logrado en el caso de haber realizado un **censo** (estudio de todos los elementos de la población).

Para el éxito de esta tarea, se han desarrollado una serie de técnicas estadísticas que establecen requerimientos a cumplirse en cada fase del proceso de investigación, empezando por la elaboración del plan de muestreo y llegando a la metodología para el análisis de la información recopilada.

El muestreo en la investigación científica y en la recolección de informaciones tiene dos ingredientes básicos:

- a. Los conceptos teóricos que establecen los principios estadísticos y matemáticos que orientan el proceso de obtención de la información en la dirección del muestreo estadístico.
- b. Los principios de aplicación práctica en un área del conocimiento o en un problema específico que establece ciertas particularidades que son fundamentales para obtener informaciones de relevancia práctica.

Ambos ingredientes pueden hacer la diferencia entre el éxito y el fracaso en un estudio por muestreo. Sí los principios de aplicación práctica dependen de la experiencia del profesional y sólo puede ser obtenida a través del ejercicio profesional, la formación de una sólida base conceptual sólo puede ser obtenida por medio del estudio de la teoría de muestreo, que se fundamenta en conceptos matemáticos y estadísticos.

Por esto, estas notas de acompañamiento se han dividido en tres capítulos. El primero contiene los conceptos estadísticos fundamentales para la comprensión de la técnica del muestreo; en el segundo capítulo se presenta la metodología para el cálculo del tamaño de muestra, considerando tres métodos de muestreo: simple aleatorio, estratificado aleatorio y sistemático. En el capítulo tres se muestran algunos ejemplos prácticos, para que el estudiante analice la correcta interpretación de resultados de investigaciones que involucran muestreo estadístico.

CAPITULO I

CONCEPTOS ESTADÍSTICOS FUNDAMENTALES

1. Individuo o Unidad Estadística

Todo estudio estadístico se hace sobre un *individuo*, que es el objeto de observación. Una unidad estadística es la entidad sobre la que se quiere obtener las informaciones para ser analizadas. Por ejemplo: un tallo de caña, una parcela con árboles, un trabajador, una llanta, un saco con azúcar, etc.

2. Población y población objetivo

2.1 Población

Es una colección completa de individuos, plantas, animales o cosas que tienen por lo menos una característica en común.

2.2 Población biológica

Una población biológica se refiere a un conjunto de individuos de una misma especie ubicada en un espacio y tiempo determinado. La población biológica se caracteriza porque sus unidades están distribuidas espacialmente en un área o en un volumen, generalmente se desconoce el tamaño de la población y la información básica disponible es a lo sumo un mapa o una fotografía aérea de la región estudiada.

El comportamiento espacial de una población biológica se analiza por la dispersión espacial que es propia de cada especie (hábitat, patrón de reproducción, etc.) y su variación refleja cambios en la densidad.

Ejemplos:

- a) Total de tallos de caña infectados por la bacteria *Leifsonia xyli* Subs. *xyli*, provocadora de la enfermedad del Raquitismo de las Socas, en una hectárea plantada con una variedad X de caña de azúcar.
- b) Área cultivada con caña en el departamento de Escuintla en la zafra 2005 – 2006.
- c) Total de técnicos capacitados durante un año sobre manejo integrado de plagas en caña.
- d) Población estimada de adultos por especie de rata cañera en la finca “San Fernando”, Escuintla.
- e) Total de parcelas (de 1000 m²) de *Pinus maximinoii* en un bosque comunal.

3. Población objetivo

Antes de iniciar el muestreo se debe definir cuidadosamente la población que va a ser muestreada. La definición debe contener una descripción de los elementos que serán incluidos y una especificación de las mediciones que se van a considerar.

En el caso de muestreo de insectos, se debe definir el estadio de la plaga y la etapa fenológica del cultivo. Por ejemplo, en chinche salivosa, se puede estar interesado en contar el número de huevos, para lo cual se toman muestras de suelo que luego son sometidas al proceso de extracción. Este muestreo se recomienda realizar inmediatamente después del corte, debido a que en ese momento es más fácil tomar las muestras, por la ausencia de caña parada.

4. Unidad de Muestreo

Las unidades de muestreo son colecciones no traslapadas de elementos de la población que cubren la población completa.

Las unidades de muestreo deben ser uniformes y estables, fácilmente identificables y deben suministrar información lo más precisa posible. Adicionalmente, cada unidad de muestreo debe tener igual probabilidad de selección, en algunos casos es deseable que permita una fácil conversión a unidades de área y debe tener una magnitud proporcional al objeto del estudio.

En el caso de poblaciones biológicas, una unidad de muestreo puede ser: plantas o partes de ella, unidades de área, surcos, trampas, etc. Ejemplos:

- a) Una macolla (grupo con un promedio de 10 tallos de caña)
- b) Cuatro macollas consecutivas,
- c) Un metro lineal de caña,
- d) El área delimitada por un marco de hierro de $30 \times 30 \times 6$ cms,
- e) El volumen que ocupa el suelo en un cilindro de 8 cm de diámetro y 10 cm de altura,
- f) Una trampa de plástico color amarillo y de forma cuadrada con dimensiones de 0.5 m por lado.

En estas unidades de muestreo se tomarán datos referentes al número de individuos presentes (número de ninfas por tallo, número de adultos por tallo, total de huevecillos) o del daño que éstos han causado (por ejemplo, total de entrenudos dañados por tallo, total de tallos dañados).

En algunas oportunidades, la unidad de muestreo más adecuada por la calidad de información que suministra, no es la mejor desde el punto de vista práctico, bien sea por el tiempo que consume la toma de la muestra, por la dificultad que implique, costo o cualquier otra razón. La solución puede ser elegir otra variable de análisis, altamente relacionada con la original, que pueda ser evaluada gracias a unidades de muestreo más cómodas de manejar. Este proceso recibe el nombre de *Muestreo Indirecto*. Duque (1996) cita el caso de la evaluación del ataque producido por *Apion godmani* Wagner en el vivero internacional de Apion en el Centro de Investigación de Agricultura Tropical de Colombia – CIAT – (1986). Para evaluar este daño se estima el porcentaje de semillas con daño. Debido al excesivo tiempo y manipulación que exigiría, se pensó en evaluar el porcentaje de vainas con daño y se estudió la relación entre las dos variables de análisis, obteniéndose buenos resultados.

Nota:

- a) Si cada unidad de muestreo contiene uno y solamente un elemento de la población, entonces una unidad de muestreo y un elemento de la población son idénticos. En el caso de encuestas de opinión, una unidad muestral puede ser una persona o un hogar.
- b) Las unidades de muestreo deben cambiar según lo haga el proceso biológico o las necesidades: si en un cultivo en alguna edad fisiológica son nocivos algunos estados y otros no, puede definirse para cada uno de ellos una unidad de muestreo propia. Por ejemplo, evaluar hojas para buscar estados inmaduros y pases de red para monitoreo de adultos, cada uno en su momento oportuno.

5. Marco

Es una lista de unidades de muestreo. Por ejemplo, si se desea estimar el rendimiento de un cultivo en una determinada localidad, puede involucrar el muestreo de una lista de productores a ser entrevistados y una lista de parcelas para ser medidas.

6. Muestra

Es una colección de unidades muestrales seleccionadas de un marco o de varios marcos. Puede decirse también que **muestra** es una porción representativa de la población. Son requisitos deseables de una buena muestra:

- a) Representatividad, que se garantiza con la selección adecuada del tipo de muestreo, y
- b) Confiabilidad (está relacionada con el tamaño de la muestra).

Ejemplos:

Plaga	Unidad de muestreo	Tamaño de la muestra (No. unidades muestrales)
Barrenador (en la cosecha)	1 caña	20 /ha
Barrenador (en la fase de pelillo)	5 ó 10 metros lineales de plantas sobre el surco	4 por lote ó 5 por hectárea
Plagas de la raíz	30 × 30 cm de área y 30 cm de profundidad	2 por hectárea
Rata cañera (monitoreo)	1 trampa tipo jaula o guillotina	2 por hectárea (para determinar el porcentaje de captura)
Chinche salivosa (adultos)	1 bolsa plástica amarilla de alta densidad, de 0.6 × 0.6 m y 3 milésimas de pulgada de grosor.	2 trampas por hectárea

El tamaño de la muestra depende del **nivel de precisión** deseado, 10% y 15% en investigación, o bien 20% o hasta 25% en toma de decisiones. El tamaño de la muestra, o dicho en otras palabras, el número de unidades de muestreo que han de constituir la muestra está determinado por la variación existente entre las unidades y por el costo implícito en la disminución de esta variación al mínimo.

Una ecuación de tipo general que puede ser aplicada sin considerar la distribución espacial de la población biológica, con la condición que el hábitat sea homogéneo, es la siguiente:

$$n = \left(\frac{s}{E \bar{x}} \right)^2$$

En la cual:

n = Número de muestra a tomar.

s = Desviación estándar obtenida con los datos de las muestras preliminares

$\bar{x}$ = Media estimada con los datos de las muestras preliminares

E = Nivel de precisión requerido.

Nota

Cuanto más acertado se desea el muestreo, mayor será el número de muestras a tomar, por lo que a menudo se hace necesario un compromiso entre precisión y capacidad de trabajo, compromiso para el cual no existen reglas salvo la de tratar de tomar el mayor número de muestras de acuerdo con la disponibilidad de recursos.

7. Censo

En ocasiones resulta posible estudiar cada uno de los elementos que componen la población, realizándose lo que se denomina un censo, es decir, el estudio de todos los elementos que componen la población. Ejemplo: IV Censo Agropecuario Nacional realizado en el 2003.

8. Parámetro

Un parámetro es un valor fijo (no aleatorio) que caracteriza a una población en particular. En general, un parámetro es una cantidad desconocida y rara vez se puede determinar exactamente su valor, por la dificultad práctica de observar todas las unidades de una población. Por este motivo, se trata de estimar el valor de los parámetros desconocidos a través del empleo de muestras. Los parámetros se representan con letras del alfabeto griego. Por ejemplo: μ = media, σ = desviación estándar, σ^2 = varianza.

9. Estimador o estadístico

Las cantidades usadas para describir una muestra se denominan estimadores o estadísticos muestrales y se representa con letras del alfabeto arábigo. Por ejemplo:

➤ Media de la muestra: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

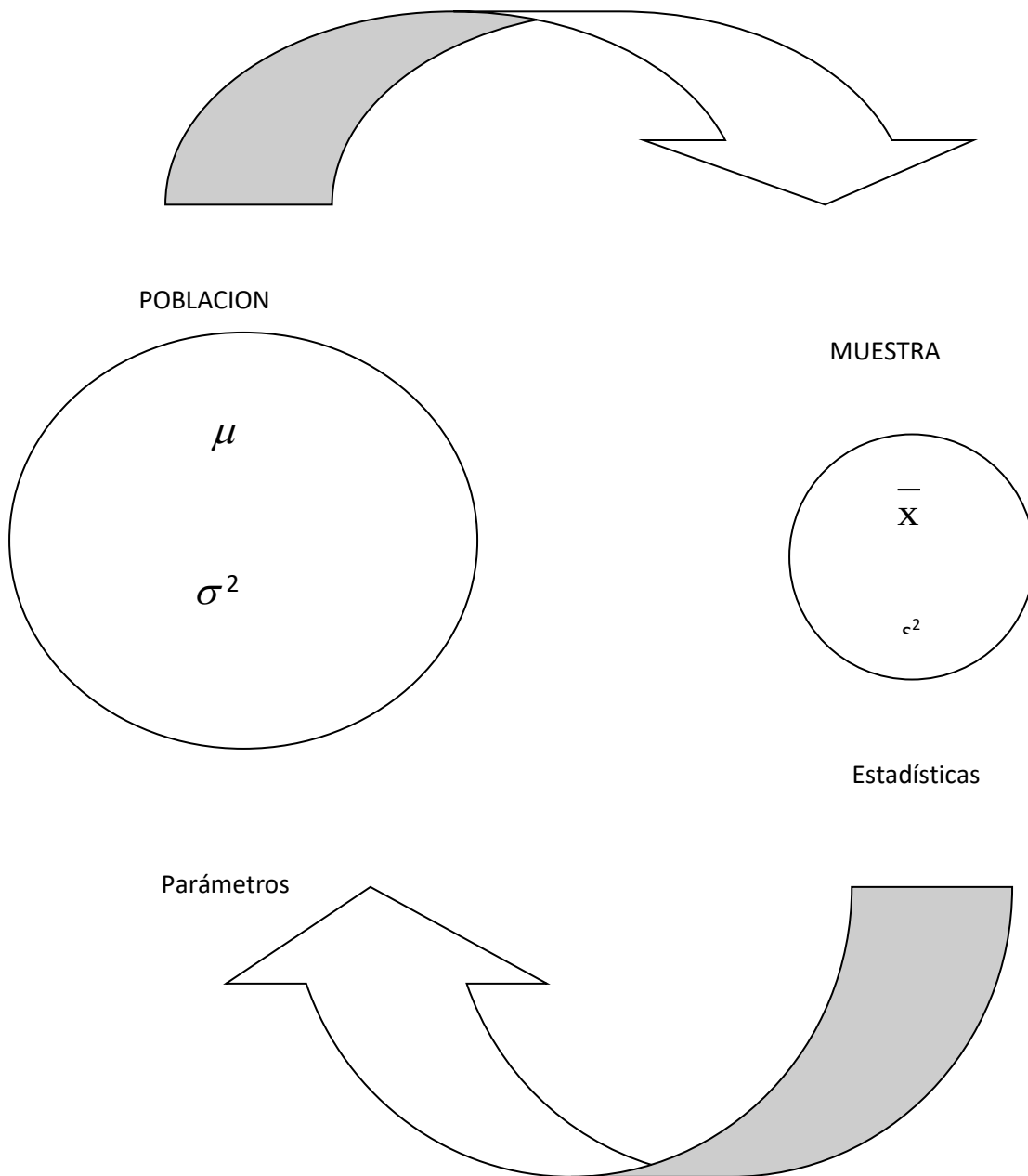
➤ Varianza de la muestra: $s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i \right)^2}{n} \right]$

➤ Menor valor de la muestra: $X_{(1)} = \text{mín} (X_1, X_2, \dots, X_n)$

➤ Mayor valor de la muestra: $X_{(n)} = \text{máx} (X_1, X_2, \dots, X_n)$

Uno de los principales problemas que se presentan en la estadística es el de hacer afirmaciones sobre los parámetros poblacionales (generalmente desconocidos), como por ejemplo, saber cuál es el tiempo necesario para que el organismo humano pueda degradar cierto compuesto químico, cuál es la producción total de granos de un determinado país en un determinado año, cuál es la altura media de la población masculina guatemalteca. Y para responder a estas preguntas, muchas veces tenemos que hacer uso del proceso de muestreo, que consiste en estudiar apenas una fracción de la población (una muestra) y a partir de ésta, hacer inferencias sobre la población.

Esquemáticamente se tiene:

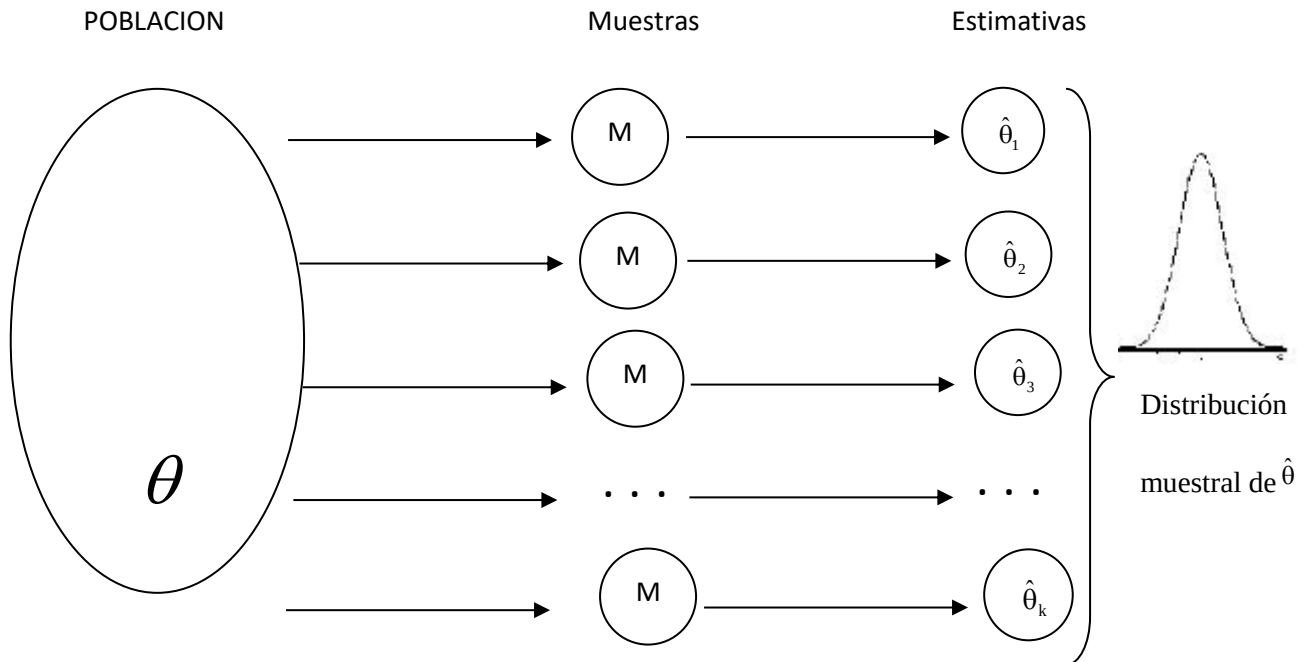


Para que el proceso anteriormente descrito sea confiable, es necesario que la muestra utilizada sea representativa de la población, y para eso, ella debe ser retirada según determinadas técnicas de muestreo.

Para hacer inferencias sobre los parámetros poblacionales de esta muestra, es necesario el conocimiento de las relaciones existentes entre las estimativas obtenidas y los valores de los parámetros poblacionales, o sea, es necesario conocer la distribución muestral del estimador utilizado, para que se pueda hacer una inferencia segura sobre un parámetro cualquiera.

10. Distribuciones de muestreo

Vamos a suponer que tenemos una población finita de tamaño N , de la cual estamos interesados en estudiar el parámetro θ (puede ser la media aritmética, total, proporción, etc.), y para ello, podemos extraer de manera aleatoria todas las k muestras (M) posibles de tamaño n , tal como se ilustra en la siguiente figura:



Podemos observar, que cada una de las muestras nos proporcionará un valor estimado del parámetro θ , el cual varía de muestra a muestra. Esta variabilidad en las estadísticas de muestras (valores estimados) proviene de un **ERROR DE MUESTREO**, debido al azar, es decir, hay diferencias entre cada muestra y la población y entre las diversas muestras, debido únicamente a las características de los individuos que decidimos seleccionar para formar las muestras.

La distribución de todos los valores que puede asumir una estadística, calculados a partir de muestras del mismo tamaño, seleccionadas en forma aleatoria de la misma población se llama: **DISTRIBUCIÓN DE MUESTRAL DE ESA ESTADÍSTICA**.

IMPORTANTE

Cualquier distribución de probabilidad (y por lo tanto, cualquier distribución de muestreo) puede ser descrita parcialmente por su media y su desviación estándar.

Distribución de muestreo de la media

Definición: Una distribución de probabilidad de todas las medias posibles de las muestras es una distribución de muestreo de la media.

Distribución de muestral de $\bar{x}$ a partir de poblaciones que siguen una distribución normal.

La distribución de muestreo de la media muestral $\bar{x}$ tiene algunas propiedades únicas. Si se obtiene una muestra aleatoria de n observaciones de una población que sigue aproximadamente una distribución normal con media μ y desviación estándar σ , la distribución de muestreo de $\bar{x}$ será normal y tendrá una media $\mu_{\bar{x}}$ (la misma media de la población muestreada) y una desviación estándar igual a $\sigma/\sqrt{n}$.

Notación: $\bar{x} \sim N(\mu, \sigma/\sqrt{n})$

La desviación estándar de la distribución de las medias de las muestras mide el grado hasta el cual esperamos que varíen las diferentes muestras debido al error de muestreo. La desviación estándar de una estadística de muestra se le conoce como: **ERROR ESTÁNDAR DE LA ESTADÍSTICA**.

Multiplicador de población finita para el error estándar de la media ($\sigma_{\bar{x}}$)

Población infinita: es aquella población en la que es teóricamente imposible observar (o medir) todos los elementos que la conforman.

Población finita: es la población que tiene un tamaño establecido o limitado. Por ejemplo: empleados de una fábrica, estudiantes de una Universidad, clientes de una agencia bancaria, etc.

Cuando la población es finita y el muestreo se realiza sin remplazo, $\sigma_{\bar{x}}$ es dado por:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

Multiplicador de población finita

Nota: Se debe utilizar el multiplicador de población finita para corregir el error estándar de la media, cuando la fracción de muestreo $n/N > 0.05$.

Distribución de muestral de $\bar{x}$ a partir de poblaciones que no siguen una distribución normal.

Cuando el muestreo se efectúa a partir de una población que no sigue una distribución normal, se utiliza un teorema matemático conocido como: **TEOREMA CENTRAL DEL LIMITE**. Este teorema se cita a continuación:

Teorema Central del Límite

Dada una población de cualquier forma funcional no normal con una media μ y desviación estándar finita σ , la distribución muestral de $\bar{x}$, calculada a partir de muestras aleatorias de tamaño n de dicha población, será aproximadamente normal con media μ y una desviación estándar igual a $\sigma/\sqrt{n}$ cuando la muestra es muy grande.

El teorema central del límite permite tomar muestras a partir de poblaciones con distribución no normal y garantizar que se obtengan aproximadamente los mismos resultados que si la población tuviera una distribución normal, siempre que se tome una muestra grande.

Relación entre el tamaño de muestra y el error estándar

$\sigma_{\bar{x}}$ $\longrightarrow$ El error estándar es una medida de dispersión de las medias de las muestras alrededor de la media de la población.

$\downarrow$ Dispersión $\downarrow \sigma_{\bar{x}}$ los valores tomados por la media de la muestra tienden a agruparse más cercanamente alrededor de μ .

- ❑ Al disminuir el error estándar, el valor de cualquier media de la muestra probablemente se acercará al valor de la media de la población, en otras palabras:
- ❑ Al disminuir el error estándar, se incrementa la precisión con la que se puede usar la media muestral para estimar la media de la población.

11. Precisión

Es el alejamiento máximo que el investigador está dispuesto a permitir entre el estimador y el parámetro correspondiente.

12. Confiabilidad

Es el grado de seguridad de que la precisión se cumpla, y se mide en términos de probabilidad. Los valores más frecuentes utilizados para representar la confiabilidad son: 1.65, 1.96, 2.58, los cuales equivalen al 90%, 95% y 99% de confianza, respectivamente.

El tamaño de la muestra guarda relación con el grado de precisión, a mayor precisión (valores cercanos a 0) mayor tamaño de muestra. En la práctica se presta más atención a la precisión que a la confiabilidad. En el cuadro siguiente se muestra como va variando el tamaño de la muestra con diferentes valores de precisión y confiabilidad. Note que a medida que aumenta el nivel de confianza también aumenta el tamaño de la muestra (fijando el valor de precisión)

Precisión (%)	Nivel de confianza (%)		
	90	95	99
5	228	306	461
10	64	90	150
15	29	42	71



N = 1,500 familias.

Muestreo utilizado: Simple Aleatorio

Parámetro a estimar: Proporción

13. Tipos de muestreo

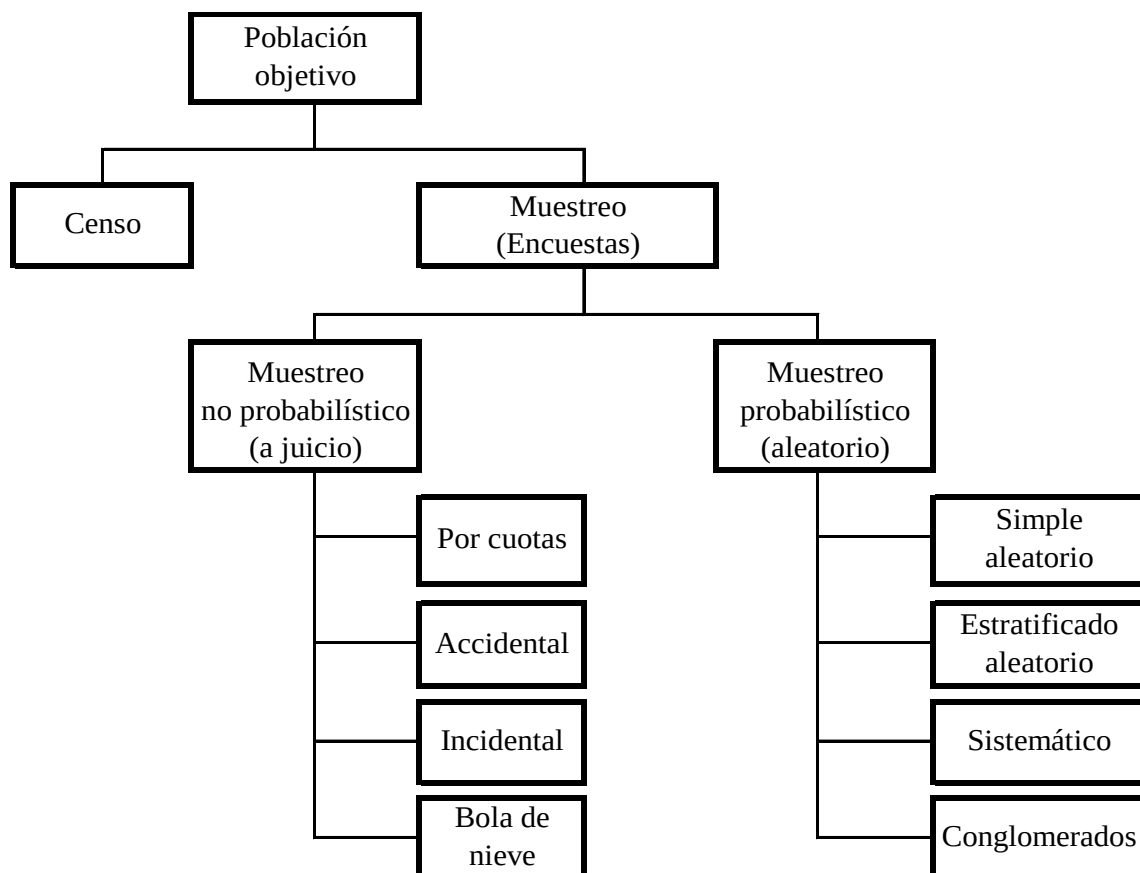
Existen dos métodos para seleccionar muestras de poblaciones: el muestreo no aleatorio (no probabilístico o a juicio) y el muestreo probabilístico (o aleatorio).

En el muestreo probabilístico, todos los elementos de la población tienen la misma oportunidad de ser seleccionados para formar parte de la muestra.

Por otra parte, en el muestreo a juicio, se emplea el conocimiento y la opinión personal para identificar aquellos elementos de la población que deben incluirse en la muestra.

Una muestra seleccionada por muestreo a juicio se basa en la experiencia de alguien con la población. Algunas veces una muestra a juicio se usa como guía o muestra tentativa para decidir cómo tomar una muestra aleatoria más adelante. Las muestras a juicio evitan el análisis estadístico necesario para hacer muestras probabilísticas, son más convenientes y pueden usarse con éxito incluso si no podemos medir su validez. Pero si un estudio emplea el muestreo a juicio y pierde un grado importante de “representatividad”, habrá adquirido conveniencia a un precio demasiado alto.

Los tipos principales de muestreo se resumen en el siguiente esquema:



13.1 Descripción de los algunos diseños de muestreo probabilístico

- **Muestreo Simple Aleatorio**

En el muestreo simple aleatorio se seleccionan muestras mediante métodos que permiten que *cada posible muestra de tamaño n tenga una igual probabilidad de ser seleccionada* y que *cada elemento de la población total tenga una oportunidad igual de ser incluido en la muestra*.

Condiciones para aplicarlo:

1. Las características de la población deben ser homogéneas con respecto a la variable de interés.
2. Se debe conocer el tamaño de la población.
3. Se debe tener un listado de todos los elementos de la población (marco lista).

Algunos casos en los que se aplica esta técnica de muestreo son presentados a continuación:

Caso 1

El muestreo simple aleatorio forma la base de la mayoría de las encuestas que se realizan en la práctica (Scheaffer et. al., 1987). El Índice de Televisión Nielsen (ITN) es el servicio que más ampliamente usado, que existe para la medición de la audiencia. Éste se basa en una muestra aleatoria de aproximadamente mil doscientos hogares que tienen conectada a la televisión un audiómetro de almacenaje instantáneo. Este medidor registra si está o no funcionando la televisión, que canal se está viendo, y cambios de canales. En una muestra aleatoria adicional de familias, cada familia anota en un diario los diferentes programas que ve. El ITN informa el número de hogares en la audiencia y el tipo de audiencia y el tiempo que ve televisión durante diversos períodos. En el Brasil, este tipo de mediciones es dado por un índice denominado IBOPE.

Caso 2

Las investigaciones de mercado frecuentemente incluyen una muestra aleatoria simple de usuarios potenciales de un producto. Los investigadores pueden desear estimar la proporción de compradores potenciales que prefieren un cierto color de automóvil o sabor de alimento.

Caso 3

Un Ingeniero forestal puede estimar el volumen de madera o proporción de árboles enfermos, seleccionando puntos geográficos en el área cubierta por el bosque y luego asociando a cada punto una parcela de forma y tamaños fijos (tal como un círculo de 10 metros de radio). Todos los árboles dentro de las parcelas seleccionadas pueden ser estudiados, pero, nuevamente, el diseño básico es una muestra simple aleatoria.

- **Muestreo Sistemático**

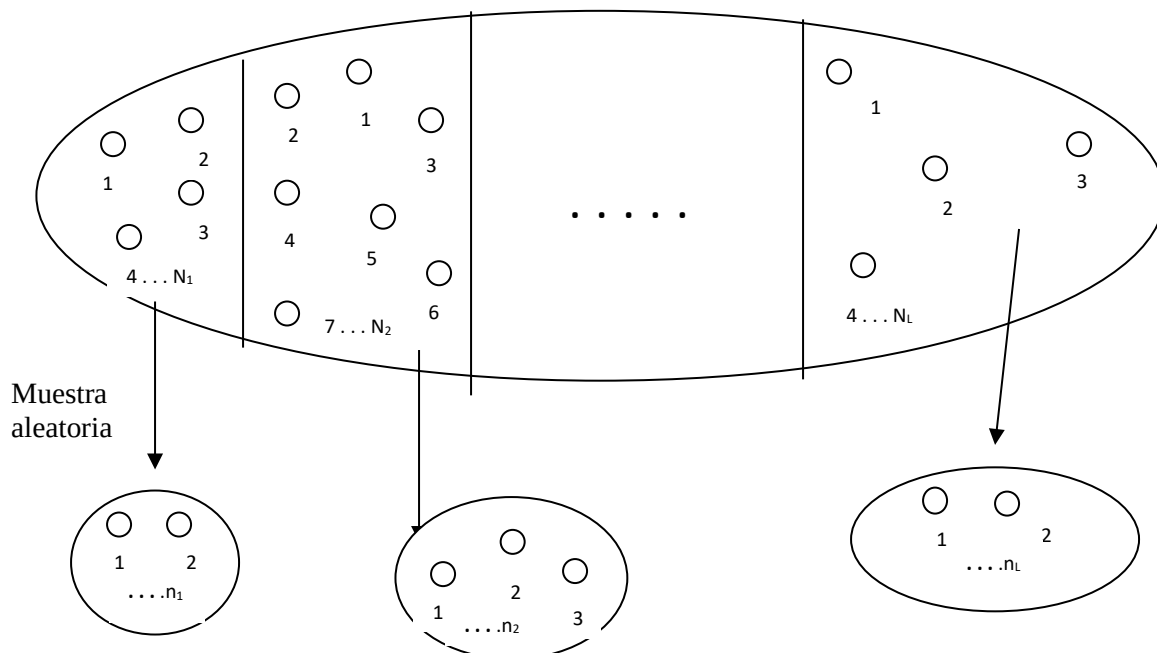
Una muestra obtenida al seleccionar aleatoriamente un elemento de los primeros k elementos en el marco y después cada k -ésimo elemento se denomina muestra sistemática de 1-en- k . En el muestreo sistemático, los elementos son seleccionados de la población dentro de un intervalo uniforme que se mide con respecto al tiempo, al orden, o al espacio. Este tipo de muestreo presenta la ventaja de ser más fácil de llevar a cabo en el campo, y por lo tanto, a diferencia de las muestras aleatorias simples y estratificadas, está menos expuesto a los errores de selección que cometen los investigadores de campo.

La idea básica del muestreo sistemático es como sigue: supóngase que una muestra de n nombres será seleccionada de una larga lista. Una manera sencilla de hacer esta selección es elegir un intervalo apropiado y seleccionar los nombres a intervalos iguales a lo largo de la lista. De este modo, cada décimo nombre podría ser seleccionado, por ejemplo.

Los investigadores de mercado y los encuestadores, quienes muestrean personas en movimiento, frecuentemente emplean un diseño sistemático. A cada vigésimo cliente en un mostrador de pago se le puede preguntar su opinión sobre el sabor, color, o textura de un producto alimenticio. A cada décima persona que aborda un autobús se le puede pedir que llene un cuestionario acerca del servicio del autobús. Cada centésimo automóvil que entra en un parque de diversiones puede ser detenido y al conductor preguntarle acerca de los sistemas de publicidad del parque o sobre el precio de los boletos. En el área forestal, pueden muestrearse sistemáticamente parcelas de terreno para estimar la proporción de árboles enfermos o pueden muestrear sistemáticamente los mismos árboles para estudiar patrones de crecimiento. El principal problema con el muestreo sistemático, es que si la población tiene un comportamiento cíclico, muestra puede ser poco representativa o existe el riesgo de introducir sesgos en la muestra.

- **Muestreo Aleatorio Estratificado**

Una muestra aleatoria estratificada es la obtenida mediante la separación de los elementos de la población en grupos relativamente homogéneos, que no presenten traslapes, llamados estratos, y la selección posterior de una muestra aleatoria simple. La estratificación pretende reunir en cada estrato a unidades homogéneas entre sí y heterogéneas en relación con los otros estratos. El muestreo estratificado resulta apropiado cuando la población ya está dividida en grupos de diferentes tamaños y se desea tomar en cuenta este hecho.



Notación:

L = número de estratos

N_i = número de unidades muestrales en el estrato i

N = número de unidades muestrales en la población (o sea $N_1 + N_2 + \dots + N_L$)

n_i = tamaño de la muestra en el i -ésimo estrato

n = tamaño de la muestra

La ventaja de las muestras estratificadas es que, cuando se diseñan adecuadamente, reflejan de manera más precisa las características de la población de la cual fueron elegidas, en comparación con otro tipo de muestras. ¿Cómo seleccionar una muestra aleatoria estratificada?

El primer paso en la selección de una muestra aleatoria estratificada es especificar claramente los estratos; así, cada unidad muestral se ubica en el estrato apropiado. Por ejemplo, la división se puede realizar con base en la topografía, los horizontes del suelo, cambios de color en el suelo, el crecimiento irregular de las plantas, etc. Esto garantiza que los puntos de muestreo se encuentren repartidos más uniformemente en toda la zona, en función al tamaño del estrato y permite además conocer de forma independiente las características particulares de cada estrato. El tamaño de la muestra puede ser proporcional o no proporcional, con relación al tamaño del estrato de la cual fue extraída.

Algunos casos en los que se aplica esta técnica de muestreo son presentados a continuación:

Caso 1

Según Scheaffer et. al. (1987), la mayoría de las grandes encuestas tiene incorporado cierto grado de estratificación en el diseño. Como ejemplo cita la Encuesta de Establecimiento (EE) realizada por la Oficina de Estadísticas de Trabajo de los Estados Unidos. Esta encuesta reúne información relativa a horas de trabajo y ganancias para establecimientos no agrícolas de los Estados Unidos. Los establecimientos son estratificados de acuerdo con el tipo de industria y tamaño, primordialmente para homogeneidad de las mediciones, pero también para aportar estimaciones para diversos tipos de industrias. Por ejemplo, se proporciona información para categorías de industrias, tales como minería, construcción, manufactura, transporte y finanzas, seguros y bienes raíces.

Caso 2

Una investigadora desea estimar el peso promedio de 90 ratas (50 machos y 40 hembras) que han sido alimentados con cierta dieta. Las ratas fueron separadas por sexo; por lo que el uso de muestreo aleatorio estratificado con dos estratos le pareció ser el más apropiado.

Caso 3

Lauretin y Pereira (2002) estudiaron el patrón de distribución y muestreo de estados inmaduros de la mosca blanca en el cultivo de ajonjolí. Para la ejecución del muestreo dividieron las plantas en tres estratos: apical, medio y basal. El estrato apical se ubicó entre 100 y 115 cm sobre el suelo, el estrato medio entre 55 y 70 cm y el estrato basal entre 10 y 25 cm. De cada uno de ellos se tomaron 5 hojas y se realizó el conteo de huevos y ninfas con la ayuda de lupas estereoscópicas.

- **Muestreo Aleatorio por Conglomerados**

Los métodos presentados hasta ahora están pensados para seleccionar directamente los elementos de la población, es decir, que las unidades muestrales son los elementos de la población. En el muestreo por conglomerados la unidad muestral es un grupo de elementos de la población que forman una unidad, a la que se denomina conglomerado. Las unidades hospitalarias, los departamentos universitarios, una caja de determinado producto, un naranjo, etc, son conglomerados naturales. En otras ocasiones se pueden utilizar conglomerados no naturales como, por ejemplo, las urnas electorales. Cuando los conglomerados son áreas geográficas suele hablarse de "muestreo por áreas". Un automóvil forma un buen conglomerado de cuatro llantas para estudio de uso y seguridad de llantas. Un naranjo forma un conglomerado de naranjas para la investigación de infestación por insectos. Una parcela en el bosque contiene un conglomerado de árboles para la estimación de volúmenes de madera o proporción de árboles enfermos.

El muestreo por conglomerados consiste en seleccionar aleatoriamente un cierto número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y en investigar después todos los elementos pertenecientes a los conglomerados elegidos.

*** **

Para finalizar con esta exposición de los métodos de muestreo probabilístico es necesario comentar que ante lo compleja que puede llegar a ser la situación real de muestreo con la que nos enfrentemos es muy común emplear lo que se denomina muestreo polietápico. Este tipo de muestreo se caracteriza por operar en sucesivas etapas, empleando en cada una de ellas el método de muestreo probabilística más adecuado.

Ejemplo:

Se desea estimar el número de granos por vaina de frijol. En la primera etapa se eligen las plantas al azar y en cada planta se eligen al azar, nuevamente, las vainas o unidades de muestreo.

*** **

13.2 Tipos de muestreo No Probabilísticos

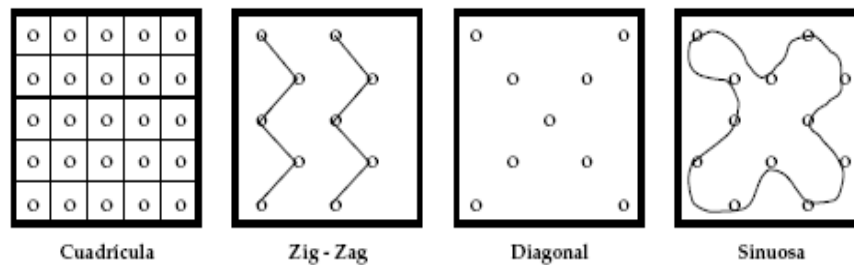
- ❑ **Muestreo por Cuotas:** este tipo de muestreo se fundamenta generalmente sobre la base de un buen conocimiento de los estratos de la población y/o individuos más “representativos” o “adecuados” para los fines de la investigación. Mantiene, por tanto, semejanzas con el muestreo aleatorio estratificado, pero no tiene el carácter de aleatoriedad de éste. En este tipo de muestreo se fijan “cuotas” que consisten en un número de individuos que reúnen unas determinadas condiciones, por ejemplo: 20 individuos de 25 a 40 años, de sexo femenino y residentes en una determinada ciudad. Una vez determinada la cuota se eligen los primeros que se encuentren que cumplan esas características. Este método se utiliza mucho en las encuestas de opinión.
- ❑ **Muestreo Accidental:** es aquel en el cual los individuos que formarán la muestra se obtienen sin ningún plan preconcebido; los individuos seleccionados resultan producto de circunstancias fortuitas. Si entrevistamos a los primeros 50 transeúntes que pasan por cierta calle, estaremos en presencia de una muestra accidental; los datos obtenidos podrán o no representar al universo en estudio. El investigador no puede saber hasta qué punto sus resultados podrán proyectarse, con confiabilidad, hacia el conjunto más amplio que desea conocer.
- ❑ **Muestreo Incidental:** se trata de un proceso en el que el investigador selecciona directa e intencionalmente a los individuos de la población. El caso más frecuente de este procedimiento es aquel en que se utiliza como muestra a individuos a los que se tiene fácil acceso.

- ❑ **Muestreo tipo “snow ball” (bola de nieve):** en este tipo de muestreo, se localiza a algunos individuos, los cuales conducen a otros, y éstos a otros, y así hasta conseguir una muestra suficiente. Este tipo de muestreo se emplea frecuentemente cuando se hacen estudios con poblaciones “marginales”, delincuentes, sectas, determinados tipos de enfermos, etc.

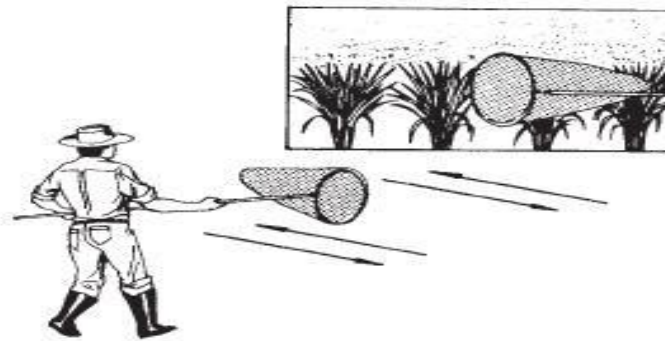
14. Toma de la muestra

Para quienes están acostumbrados a hacer muestreos no es desconocido el hecho de que diversas técnicas para tomar muestras pueden conducir a diferentes resultados y es frecuente que las técnicas más adecuadas por su precisión pueden tener dificultades en ser adoptadas. Para la solución de este problema se sugiere hacer una estimación de la capacidad de estas técnicas para evaluar la variable de análisis y de esta manera poder definir factores de corrección o calibración para ellas.

Por ejemplo, en el caso de muestreo de suelos, el recorrido a efectuar en la toma de muestra es de acuerdo a alguno de los siguientes esquemas:



En el caso de muestreo de insectos en el follaje, se puede considerar los anteriores recorridos. Para la toma de la muestra, se hacen barridos utilizando una red entomológica. El barrido es un pase de red en la dirección que se toma la muestra y describiendo un ángulo de 180° frente a cuerpo de quien toma la muestra. Después de cada paso de red, el muestreador da un paso adelante. El CIAT (Colombia) recomienda para el muestreo de artrópodos en el cultivo del arroz, 100 pases de red distribuidos en 5 sitios o áreas, es decir, en cada uno de éstos se hacen 20 pases.



Toma de la muestra de artrópodos en el cultivo del arroz.

Fuente: Nociones Básicas del MIP en Arroz. CIAT, Colombia

15. Errores en un estudio por muestreo

Aunque se empleen métodos de muestreo probabilísticos, las investigaciones están sujetas a errores potenciales. A continuación se describen los tipos de errores de investigación:

1. Error de muestreo

El error de muestreo aparece como consecuencia de la utilización de una parte de la población para estimar características de todo el conjunto poblacional. En general, a medida que aumenta el tamaño de la muestra disminuye el error de muestreo, aunque también influyen en él las características del diseño y la naturaleza de la variable investigada.

2. Errores de no muestreo

Se trata de errores que no tienen que ver con el hecho de seleccionar una muestra. Puede ser consecuencia de la naturaleza del diseño del estudio y de las imperfecciones en su ejecución. Básicamente se tienen tres tipos de errores de no muestreo:

2.1 Error de cobertura o sesgo en la selección.

2.2 Error por falta de respuesta o sesgo por falta de respuesta.

2.3 Error de medición.

Un buen proyecto de estudio de investigación intenta reducir o minimizar esos diversos errores, frecuentemente a un costo considerable.

CAPITULO II

PROCEDIMIENTO PARA EL CÁLCULO DEL TAMAÑO DE LA MUESTRA

2.1 MUESTREO SIMPLE ALEATORIO (MSA)

2.1.1 Definición

En el muestreo simple aleatorio se seleccionan muestras mediante métodos que permiten que *cada posible muestra de tamaño n tenga una igual probabilidad de ser seleccionada y que cada elemento de la población total tenga una oportunidad igual de ser incluido en la muestra.*

2.1.2 Ventajas e inconvenientes del uso del MSA

El muestreo simple aleatorio es sencillo y muy eficiente cuando se aplica en poblaciones con unidades de muestreo homogéneas en cuanto al parámetro poblacional que se desea estimar (volumen total, área basal, número de árboles por hectárea, etc.)

Sin embargo, es poco práctico cuando se trata de inventariar bosques tropicales, debido a que¹:

- a) Muchas veces las parcelas que constituyen la muestra quedan en sitios inaccesibles o alejados del resto de las parcelas, lo cual tiene repercusión en los costos.
- b) No proporciona datos confiables acerca de la configuración y topografía del bosque, que si bien no son la información más importante que se genera con un inventario, son por lo general datos complementarios de mucha utilidad para la planificación del manejo forestal.
- c) Hay poca seguridad sobre el establecimiento exacto de las parcelas en el bosque.

2.1.3 Determinación del tamaño de la muestra

Pasos para determinar el tamaño de la muestra

1. Obtener una muestra piloto (puede tomarse un 10% de la población). La muestra preliminar o piloto, es una muestra que antecede a la definitiva, cuya selección se hace de acuerdo a los lineamientos que marca el diseño de muestreo que se utilizará en el estudio definitivamente.

La muestra preliminar juega un papel importante en el diseño de un estudio por muestreo, ya que será la fuente de información inmediata para:

- i. Tener una primera aproximación de los costos que se involucran en el estudio.
- ii. Tener una primera aproximación del tiempo que se llevará en la realización del estudio.
- iii. Estimar los parámetros involucrados en la determinación del tamaño de la muestra, usualmente la media y la varianza.
- iv. Probar la factibilidad de: los métodos de selección de las unidades muestrales, la medición de las variables y otros aspectos prácticos.
- v. Probar la factibilidad del cuestionario, como medio para la captación de la información.

¹ Inventarios forestales para bosques latifoliados en América Central/ eds. Lorena Orozco, Cecilia Brumér. –Turrialba, C.R. : CATIE, 2002, 264 p.

Algunos autores sugieren que la muestra preliminar podrá considerarse como parte de la muestra definitiva, solamente cuando los métodos de selección, medición, incluyendo el cuestionario, no hayan sufrido cambios o modificaciones severas.

2. Numerar los N elementos de la población.
3. Sortear sin reposición, n números comprendidos de 1 a N. Este proceso de sorteo puede ser realizado a través de cualquiera de los siguientes procedimientos:

3.1 Función Ran# (disponible en las calculadoras).

3.2 Con la tabla de número aleatorios.

- a) Enumerar los N elementos de la población de la siguiente forma:

Si el número de dígitos de N es igual a 2, utilice 01, 02, 03, . . . , N

Si el número de dígitos de N es igual a 3, utilice 001, 002, 003, . . . , N

- b) Iniciar en un punto cualquier de la tabla de números aleatorios.
- c) Leer los números sistemáticamente siguiendo las columnas de la tabla. Comenzando en la columna de la izquierda.
- d) Si un número ya leído surge nuevamente o si aparece un número que no está presente en la enumeración de los elementos de la población, ignorarlo y seguir con el próximo.
- e) Continuar hasta que n unidades sean seleccionadas.

- 3.3 El programa Excel también ofrece una alternativa para generar números aleatorios:

Posicionese en una celda y digite (o busque en la función f_x) la siguiente fórmula =ALEATORIO.ENTRE(1,100), para el caso que tengamos 100 datos. Luego copie esta fórmula en la cantidad de celdas necesarias (tamaño de la muestra).

3.4 Extracción de papelitos numerados. Este método es sencillo, pero laborioso si la población es grande, y consiste en hacer papelitos debidamente numerados entre 1 y N. Se mezclan perfectamente en una bolsa y se extraen sin sustitución uno hasta completar n , el tamaño de la muestra.

4. Calcular el tamaño de la muestra definitiva, en este caso, para estimar la **media**, a través de la siguiente expresión (para poblaciones finitas):

$$n = \frac{N \times s^2 \times z_{(\alpha/2)}^2}{N \times d^2 + (s^2 \times z_{(\alpha/2)}^2)}$$

en que:

n = Tamaño de la muestra definitiva

N = Tamaño de la población

s^2 = Varianza obtenida con la muestra piloto

d = Precisión del estimador de interés (en este caso, la media)

α = Nivel de significancia

$z_{(\alpha/2)}$ = Valor tal que $P(|Z| < z_{(\alpha/2)}) = 1 - \alpha$, y Z es una variable con distribución normal estandarizada.

Ejemplo:

Juan Pérez, gerente general de la Cooperativa de Alimentos Milán (CAM), está interesada en tratar de comprender los hábitos de compra de los 500 miembros de la CAM, a través del estudio de algunos datos, tales como: gastos reales semanales en alimentos, número de personas en la unidad familiar, ingreso familiar anual real, etc. Por lo que utilizó un cuestionario para recolectar datos sobre los socios. Durante el mes de junio todos los miembros visitaron la cooperativa por lo menos una vez. Por tanto, existían datos sobre cada uno de los socios.

Juan espera poder utilizar este conocimiento para planificar mejor la variedad y el volumen de artículos que la CAM ofrece.

Con el propósito de conocer mejor a los socios, Juan quiere saber inicialmente el **promedio semanal de gastos en alimentos** (expresados en dólares americanos). Para ello, tomó inicialmente una muestra preliminar de 20 socios. Los resultados se presentan a continuación:

No. Identificación de la Unidad Familiar	Gasto real semanal en alimentos	No. Identificación de la Unidad Familiar	Gasto real semanal en alimentos
169	30	59	30
249	15	339	30
354	51	271	20
84	30	112	20
424	75	415	51
290	38	69	28
95	60	175	45
284	45	282	69
88	75	49	15
22	30	75	87

Con esta información determine el tamaño de muestra, utilizando una precisión del 10% y un nivel de 95% de confianza.

Cálculos:

$$N = 500 \text{ socios}$$

$$\bar{x} = \frac{1}{20} \sum_{i=1}^{20} x_i = \text{US\$ } 42.20$$

$$d = (42.20 \times 0.10) = 4.22$$

$$s^2 = 462.8 \text{ u}^2$$

$$P(Z) = (1 - \alpha) / 2 = 0.475 \quad Z \text{ (valor tabular)} = 1.96$$

$$n = \frac{500 \times 462.8 \times 1.96^2}{500 \times 4.22^2 + (462.8 \times 1.96^2)}$$

$$n = 83 \text{ socios}$$

- Para calcular el tamaño de la muestra definitiva, en este caso, para estimar el **TOTAL**, suponiendo distribución normal del estimador, con precisión y confiabilidad dadas, se utiliza el siguiente procedimiento, ilustrado con un ejemplo:

Si se considera la variable ingreso familiar anual real (expresados en dólares americanos), determine el tamaño de muestra para estimar el ingreso total. Utilice un nivel de confianza al 95% y una precisión de 15%

$$n = \frac{N^2 \times s^2 \times z^2_{(\alpha/2)}}{d^2 + (N \times z^2_{(\alpha/2)} \times s^2)}$$

Siendo:

n = Número de elementos a seleccionar en la muestra definitiva.
N = Tamaño de la población
S² = Varianza obtenida en la muestra preliminar
z_(α/2) = Valor tal que P(|Z| < z_(α/2)) = 1 - α, y Z es una variable con distribución normal estandarizada.
d = Precisión

Cálculos:

En la siguiente tabla se presentan los resultados obtenidos con la muestra piloto (o inicial)

No. Identificación de la Unidad Familiar	Ingreso familiar real	No. Identificación de la Unidad Familiar	Ingreso familiar real
179	8000	309	14000
186	8700	336	6000
485	7000	475	54000
80	2500	483	8000
104	3500	230	22000
293	6000	375	6000
74	2500	51	1500
70	6000	29	11000
347	4500	247	9000
337	4500	5	8000

$$N = 500 \text{ socios} \quad \bar{x} = \frac{192700}{20} = \text{US\$ } 9635 \quad s^2 = 129,975,026 \text{ u}^2$$

$$\hat{\tau} = N \times \bar{x} = \text{US\$ } 9635 / \text{ familia} \times 500 \text{ familias} = \text{US\$ } 4817500 \text{ (total estimado)}$$

$$d = 0.15 \times \hat{\tau} = 0.15 \times \text{US\$ } 4817500 = 722625$$

$$P(Z) = (1 - \alpha) / 2 = 0.475 \quad Z \text{ (valor tabular)} = 1.96$$

$$n = \frac{500^2 \times 129975026 \times 1.96^2}{722625^2 + (500 \times 1.96^2 \times 129975026)} = 162 \text{ socios.}$$

6. Para calcular el tamaño de la muestra definitiva para estimar una **PROPORCIÓN**, se utiliza la siguiente expresión (para poblaciones finitas*):

$$n = \frac{N \times p \times q \times z_{(\alpha/2)}^2}{(d^2 \times N) + p \times q \times z_{(\alpha/2)}^2} \quad (1)$$

en que:

- n = Tamaño de la muestra definitiva
- N = Tamaño de la población
- p = Proporción de éxito de la variable obtenida en un premuestreo.
- q = Proporción de fracaso de la variable obtenida en un premuestreo.
- d = Precisión del estimador de interés (en este caso, la proporción)
- α = Nivel de significancia
- $z_{(\alpha/2)}$ = Valor tal que $P(|Z| < z_{(\alpha/2)}) = 1 - \alpha$, e Z es una variable con distribución normal estandarizada.

* Como población finita se considera a aquella población que tenga menos de 500,000 elementos)

Ejemplo 1:

Un estudiante de Mercadotecnia está planificando realizar una investigación para determinar la proporción de hogares que tienen refrigerador; por lo que es necesario calcular el tamaño de la muestra requerida con un nivel de confianza del 95% y una precisión del 5%.² La investigación se llevará a cabo en una población de 1,500 familias.

- N = 1,500
- p = 0.5
- q = 0.5
- d = 0.05
- α = 0.05
- $z_{(\alpha/2)}$ = 1.96

$$n = \frac{1,500 \times 0.5 \times 0.5 \times 1.96^2}{(0.05^2 \times 1,500) + 0.5 \times 0.5 \times 1.96^2} = 306$$

R\ Serán necesarias 306 entrevistas.

Ejemplo 2:

Un grupo de estudiantes de la Facultad de Agronomía realizará un estudio sobre las condiciones socioeconómicas de los 400 adjudicatarios del bosque comunal de la aldea Panyebar, San Juan La Laguna, del departamento de Sololá. Y desean determinar cuál será la cantidad de cuestionarios que tendrán que elaborar, si fijan un nivel de confianza del 95% y una precisión del 15%.

Para situaciones cuando:

- a) El nivel de confianza es del 95%
- b) Los valores de p y q se desconocen y se asume 0.5 para cada uno (o sea, se está asumiendo varianza máxima) La ecuación (1) se reduce a la siguiente expresión:

² Fisher & Navarro (1997) citan que en estudios de investigación de mercados el valor usual de precisión está entre 2% y 6%.

$$n = \frac{N}{N \times d^2 + 1}$$

Para el ejemplo 2 tenemos que:

$$\begin{aligned} N &= 400 \\ d &= 0.15 \\ \alpha &= 0.05 \\ Z_{(\alpha/2)} &= 1.96 \end{aligned}$$

$$n = \frac{400}{400 \times 0.15^2 + 1} = 40$$

R\ Serán necesarias 40 entrevistas.

2.1.4 Muestreo en inventarios forestales

Los inventarios por muestreo son los más utilizados en la planificación del manejo forestal, ya que permiten determinar las características de la población (estadísticos de posición y variación) con respecto a diferentes variables, a un costo apropiado y dentro de límites de tiempo razonables.

El proceso de muestreo en inventarios forestales consiste básicamente en:

1. Definir las unidades de muestreo (generalmente parcelas de tamaño y forma específica)
2. Definir la población, el marco de muestreo (mapas, listas, esquemas, etc.) y el número total de unidades muestreo (N), en el caso de poblaciones finitas.
3. Identificar cada unidad de muestreo en la población.
4. Definir el número de unidades de muestreo que serán evaluadas, esto es, el tamaño de la muestra (n). El tamaño de la muestra depende de:
 - ☐ Tiempo disponible
 - ☐ Recursos económicos disponibles
 - ☐ Intensidad de muestreo o el grado de error de muestreo deseado (generalmente establecidos por las normas de calidad o la legislación existente)
5. De la población total, seleccionar las n unidades de muestreo que serán evaluadas.
6. Identificar en el campo cada una de las unidades de muestreo.
7. Evaluar o medir las variables previamente definidas en cada una de las unidades de muestreo seleccionadas.
8. Calcular y/o estimar la información requerida, según el diseño del inventario.
9. Preparar el informe con base en el diseño previamente elaborado.

A continuación se presentan algunas definiciones básicas para entender el procedimiento de cálculo del tamaño de muestra en el área forestal.

INTENSIDAD DE MUESTREO (i)

Se define como la proporción de unidades de muestreo seleccionadas como parte de la muestra, es decir, que i es el resultado de dividir el tamaño de la muestra n entre el tamaño de la población N , o sea, $i = n / N$.

En inventarios forestales donde las unidades de muestreo son parcelas, la intensidad de muestreo también se puede calcular utilizando el área efectiva de evaluación en hectáreas (A) y el área o tamaño de las parcelas en hectáreas (a).

Ejemplo: Si se desea inventariar un bosque de 10,000 ha y se utiliza como muestra 70 parcelas de 1 ha, la intensidad de muestreo sería:

$$i = \frac{70 \times 1 \text{ ha}}{10,000 \text{ ha}} \times 100 = 0.7\%$$

La intensidad de muestreo en este caso fue de 0.7 % del área total.

COEFICIENTE DE VARIACIÓN (CV)

Es una medida muy útil para comparar la variación entre diferentes conjuntos de datos o poblaciones. El coeficiente de variación por lo general se expresa en porcentaje, y se calcula de la siguiente forma:

$$CV(\%) = \frac{s}{\bar{x}} \times 100$$

ERROR DE MUESTREO

En inventarios forestales se conoce como error de muestreo absoluto (E) al resultado de multiplicar el error estándar y el coeficiente de confiabilidad. El valor del error de muestreo depende del nivel de significancia (α) seleccionado para determinar el coeficiente de confiabilidad.

El error de muestreo también puede expresarse en forma relativa como un porcentaje de error ($E\%$), el cual se calcula dividiendo el error de muestreo absoluto entre el valor de la estimación puntual.

$$E_{(1-\alpha)} = t_{(\alpha/2, n-1)} \times S_x$$
$$E\%_{(1-\alpha)} = \frac{E_{(1-\alpha)}}{\bar{x}} \times 100$$

El error de muestreo absoluto permite determinar los límites de confianza de la media poblacional, ya que:

$$IC_{(1-\alpha)} = \bar{x} - t_{(\alpha/2, n-1)} \times S_x \leq \mu \leq \bar{x} + t_{(\alpha/2, n-1)} \times S_x$$

Nota: El error de muestreo es diferente para cada parámetro forestal y también difiere de especie a especie.

CALCULO DEL NÚMERO DE PARCELAS

A partir de la ecuación de error de muestreo relativo se puede encontrar una ecuación que permita calcular el tamaño de muestra (para poblaciones infinitas) necesario para una precisión deseada. Veamos:

$$E\%_{(1-\alpha)} = \frac{E_{(1-\alpha)}}{\bar{x}} \times 100 \qquad E_{(1-\alpha)} = t_{(\alpha/2, n-1)} \times S_x$$

Partiendo de: , y tomando en cuenta que:

Llegamos a la siguiente expresión:

$$E\%_{(1-\alpha)} = \frac{t_{(\alpha/2, n-1)} \times S_{\bar{x}}}{\bar{X}} \times 100 \text{ , y como sabemos que } S_{\bar{x}} = \frac{s}{\sqrt{n}} \text{ , tenemos que:}$$

$$E\%_{(1-\alpha)} = \frac{t_{(\alpha/2, n-1)} \times \frac{s}{\sqrt{n}}}{\bar{X}} \times 100 \text{ .}$$

De la ecuación del coeficiente de variación: $(CV\%) = \frac{s}{\bar{X}} \times 100$ despejamos la desviación estándar (s) y la sustituimos en la última ecuación de $E\%_{(1-\alpha)}$:

$$s = \frac{\bar{X} \times (CV\%)}{100} \quad \Rightarrow \quad E\%_{(1-\alpha)} = \frac{t_{(\alpha/2, n-1)} \times \frac{\bar{X} \times (CV\%)}{100 \times \sqrt{n}}}{\bar{X}} \times 100$$

$$\Rightarrow E\%_{(1-\alpha)} = \frac{t_{(\alpha/2, n-1)} \times \bar{X} \times (CV\%)}{\bar{X} \times 100 \times \sqrt{n}} \times 100 \quad \Rightarrow \quad E\%_{(1-\alpha)} = \frac{t_{(\alpha/2, n-1)} \times (CV\%)}{\sqrt{n}}$$

$$\Rightarrow \sqrt{n} = \frac{t_{(\alpha/2, n-1)} \times (CV\%)}{E\%_{(1-\alpha)}} \quad \text{De donde se obtiene:}$$

$$n = \frac{t_{(\alpha/2, n-1)}^2 \times (CV\%)^2}{E\%_{(1-\alpha)}^2}$$

El problema normalmente, es que de antemano no se conoce el valor del coeficiente de variación. Este valor depende de la homogeneidad del bosque y del tamaño de las unidades de muestreo. La estimación del coeficiente de variación puede basarse en valores de bosque cercanos, anteriormente inventariados o mediante la realización de un pre-muestreo o muestreo piloto, aunque este último método ha demostrado ser poco práctico.

Para el caso de poblaciones finitas (N= tamaño de población conocido), la ecuación para estimar n es la siguiente:

$$n = \frac{t_{(\alpha/2, n-1)}^2 \times (CV\%)^2}{E\%_{(1-\alpha)}^2 + \frac{t_{(\alpha/2, n-1)}^2 \times (CV\%)^2}{N}}$$

Para fines prácticos, la corrección para población finita se puede eliminar, porque la diferencia en el resultado final va a ser menor de una parcela.

Para calcular el tamaño de la muestra n con las ecuaciones anteriormente descritas, primero se calcula n asumiendo un valor de $t_{(\alpha/1, n-1)}$ arbitrario; por ejemplo $t_{(0.025, 6)} = 2.570$. Si se tienen datos de un muestreo preliminar, el valor de t de Student se puede determinar utilizando los grados de libertad de ese inventario preliminar.

Una vez que se ha calculado el tamaño de la muestra, se vuelve a calcular nuevamente, pero esta vez, variando el número de grados de libertad. Este procedimiento se repite nuevamente hasta que el valor estimado de tamaño de muestra no cambie.

Ejemplo:

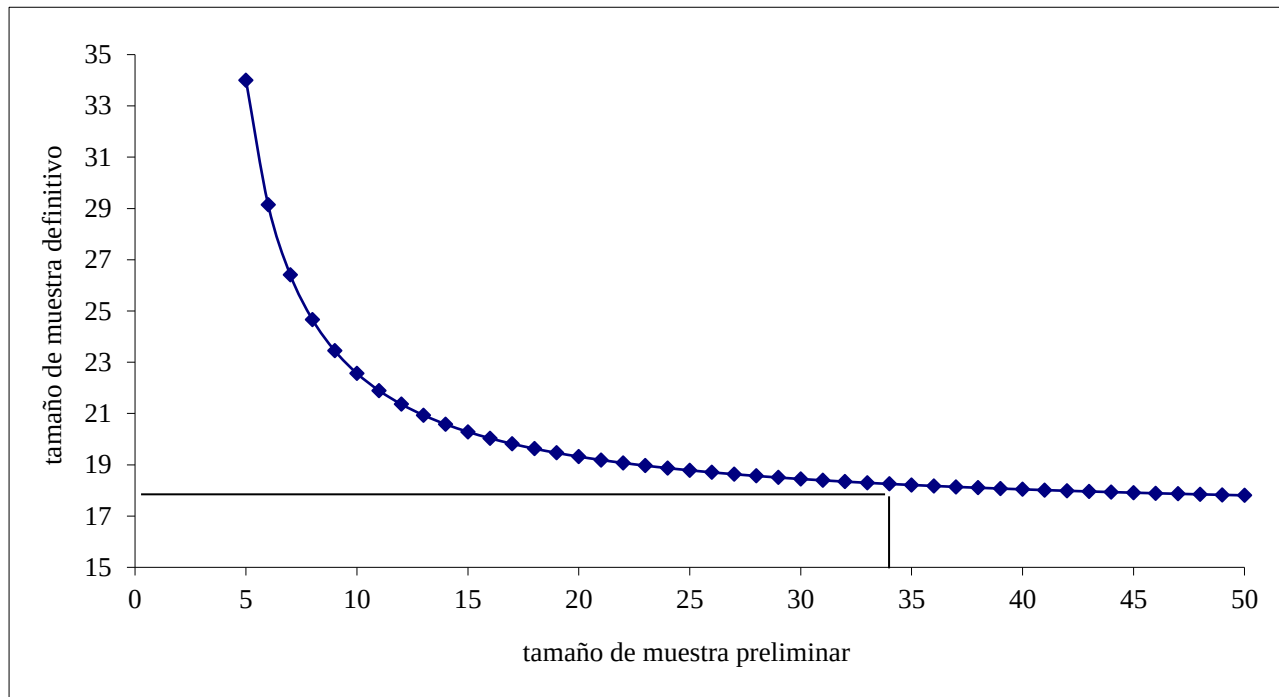
En un estudio preliminar se tomó una muestra de 10 parcelas de 1 hectárea de un bosque de 1000 ha. Los volúmenes encontrados fueron los siguientes: 64, 47, 39, 67, 54, 33, 49, 85, 72 y 90 m³/ha. Con esta información obtenga:

- El coeficiente de variación
- La media aritmética
- La desviación estándar
- La estimación del error estándar
- El error de muestreo absoluto y relativo

Encuentre el tamaño de la muestra adecuado, si desea que el error de muestreo sea de 15%, con un nivel de confianza de 95%.

tamaño de muestra preliminar	gl	valor de t	tamaño muestra definitivo	tamaño de muestra preliminar	gl	valor de t	tamaño muestra definitivo
5	4	2.776	34	28	27	2.052	19
6	5	2.571	29	29	28	2.048	19
7	6	2.447	26	30	29	2.045	18
8	7	2.365	25	31	30	2.042	18
9	8	2.306	23	32	31	2.040	18
10	9	2.262	23	33	32	2.037	18
11	10	2.228	22	34	33	2.035	18
12	11	2.201	21	35	34	2.032	18
13	12	2.179	21	36	35	2.030	18
14	13	2.160	21	37	36	2.028	18
15	14	2.145	20	38	37	2.026	18
16	15	2.131	20	39	38	2.024	18
17	16	2.120	20	40	39	2.023	18
18	17	2.110	20	41	40	2.021	18
19	18	2.101	19	42	41	2.020	18
20	19	2.093	19	43	42	2.018	18
21	20	2.086	19	44	43	2.017	18
22	21	2.080	19	45	44	2.015	18
23	22	2.074	19	46	45	2.014	18
24	23	2.069	19	47	46	2.013	18
25	24	2.064	19	48	47	2.012	18
26	25	2.060	19	49	48	2.011	18
27	26	2.056	19	50	49	2.010	18

En forma gráfica, se pueden representar los resultados del cuadro anterior, y determinar el tamaño de la muestra.



En la siguiente dirección electrónica, encontrará una calculadora que le ayudará para determinar el tamaño de una muestra cuando se estudia una proporción MSA.

2.2 MUESTREO ALEATORIO ESTRATIFICADO (MAE)

2.2.1 Definición

Si una población de tamaño N es dividida en grupos homogéneos que no presentan traslapes, llamados ESTRATOS, de tamaño N_i ($i=1, 2, \dots, k$) y luego se toma una muestra aleatoria simple de cada uno de ellos, la muestra final se llamará: MUESTRA ALEATORIA ESTRATIFICADA, la cual es la suma de todas las muestras de cada estrato.

2.2.2 Ventajas e inconvenientes

a) Ventajas

La estratificación es una técnica usual por distintas razones, entre otras:

- Produce estimaciones más precisas que las que se obtienen a partir del muestreo simple aleatorio.
- El costo por observación puede ser reducido mediante la estratificación de la población.
- Se puede obtener información de parámetros poblacionales para cada estrato de la población.
- Se simplifica el trabajo administrativo y el de control, ya que se puede usar personal específico para cada estrato.

- El tamaño de muestra será menor, si la estratificación es bien definida, en comparación con el muestreo simple aleatorio.

b) Inconvenientes

- En poblaciones con gran variabilidad, los estratos muchas veces resultan muy pequeños en extensión, por lo que no se puede hacer un muestreo de pequeña densidad dentro de cada estrato.

Si los estratos son muy diferentes en condiciones para el manejo, puede ser útil considerarlos como poblaciones aparte y hacer inventarios para cada uno, con el fin de recabar datos con suficiente confiabilidad por estrato según las variables de importancia en cada estrato.

2.2.3 ¿Cómo seleccionar una muestra aleatoria estratificada?

El primer paso en la selección de una muestra aleatoria estratificada es especificar claramente los estratos, en cuanto a su conformación y número, así, cada unidad muestral se ubica en el estrato apropiado. En general, los estratos naturales o convenientemente definidos deberán ser homogéneos internamente y heterogéneos entre ellos, con respecto a la variable bajo estudio.

Para la población así dividida, habrá que determinarse dos tamaños de muestra:

- Para la población (n), y
- Para cada estrato (n_i)

Para el caso de n , el procedimiento que será utilizado para su determinación, será aquel donde generalmente se supone que el estimador del parámetro tiene distribución normal y se establece de antemano una precisión y confiabilidad.

Resuelto el tamaño, ahora habrá que asignarlo o distribuirlo entre los estratos de tal manera que:

$$n_1 + n_2 + \dots + n_k = n$$

El tamaño de la muestra en cada estrato se puede establecer de varias formas:

- Asignación igual

Esta asignación se hace cuando los estratos son similares en cuanto a variabilidad, tamaño, así como costo por unidad, por lo tanto:

$$n_i = \frac{n}{k}, \quad i = 1, 2, \dots, k$$

siendo k = número de estratos.

- Asignación proporcional o de Bowley

En este procedimiento de fijación de la muestra por estrato, el tamaño de la muestra (n) se distribuye proporcionalmente al tamaño del estrato. En este caso, si se tiene una población de tamaño N , dividida en k estratos de tamaño N_i , en donde la suma de los $N_i = N$, se tiene que la proporción del tamaño de cada estrato (P_i) es igual a N_i / N . Para calcular el tamaño de muestra (n_i) para cada estrato, se procede a multiplicar n por P_i :

$$n_i = n \times P_i, \quad i = 1, 2, \dots, k, \text{ siendo } P_i = \frac{N_i}{N}.$$

Después que las unidades de muestreo han sido divididas en estratos, se selecciona una muestra simple aleatoria de cada estrato, de la misma manera que en la técnica de muestreo simple aleatorio.

c) Proporcional a la variación en el estrato

En este procedimiento de fijación de la muestra por estrato, también conocido como fijación óptima, el tamaño de la muestra (n) se distribuye proporcionalmente a la variación del estrato medida con un estimado de la desviación del estrato (s_i). En este caso, si se tiene una población de tamaño N dividida en k estratos de tamaño N_i , en donde la suma de los $N_i=N$, se tiene que la proporción del tamaño de cada estrato (P_i) es igual a N_i/N . Para calcular el tamaño de muestra n_i para cada estrato, se utiliza la siguiente ecuación:

$$n_i = n \times \frac{P_i \times s_i}{\sum_{i=1}^k P_i \times s_i}.$$

2.2.4 Ejemplo de cálculo del tamaño de muestra en un inventario estratificado

Un bosque natural de 300 hectáreas se ha dividido en tres estratos considerando su posición fisiográfica en: tierra alta ($A=80$ ha), terraza de declive ($B=100$ ha) y terraza baja ($C=120$ ha). Para conocer su variabilidad y calcular el tamaño de la muestra (n), se realizó un muestreo piloto en cada estrato. Se obtuvieron los siguientes datos de volumen total (en m^3), por parcela, extrapolados a hectárea.

Parcela	Estrato		
	A	B	C
1	16	7	20
2	0	10	10
3	24	22	8
4	8	6	22
5	39	11	18
Suma	87	56	78
Promedio	17.4	11.2	15.6
Desviación estándar	15.027	6.380	6.229

La proporción que guardan los estratos con respecto al área total sería:

Estrato A: $P_1 = \frac{a_1}{A} = \frac{80}{300} = 0.27$

Estrato C: $P_3 = \frac{a_3}{A} = \frac{120}{300} = 0.40$

Estrato B: $P_2 = \frac{a_2}{A} = \frac{100}{300} = 0.33$

Ahora calculemos la desviación estándar común, $Pse = \sum_{i=1}^k P_i \times s_i$

Estrato	S_i	P_i	$S_i * P_i$
A	15.027	0.27	4.01
B	6.38	0.33	2.13
C	6.229	0.40	2.49
Pse =			8.63

Como el error permisibles está expresado en porcentaje, también el Pse debe convertirse en porcentaje. Para ello, primero se obtiene el volumen medio (V), utilizando la ecuación de la media ponderada:

$$\bar{X}_p = \frac{\sum_{i=1}^k \bar{X}_i \times N_i}{\sum_{i=1}^k N_i} = \frac{(17.4 \times 80) + (11.2 \times 100) + (15.6 \times 120)}{(80 + 100 + 120)} = \frac{4384 \text{ m}^3}{300 \text{ ha}} = 14.61 \text{ m}^3/\text{ha}$$

De la anterior ecuación, se obtiene la desviación estándar de la población estratificada (Pse %):

$$Pse\% = \frac{Pse \times 100}{V} = \frac{8.63 \times 100}{14.61} = 59.07\%$$

Para poder obtener el tamaño de muestra (n) definitivo se utiliza la siguiente ecuación:

$$n = \frac{t_{(\alpha/2)}^2 \times (Pse\%)^2}{E\%^2 + \frac{t_{(\alpha/2)}^2 \times (Pse\%)^2}{N}}$$

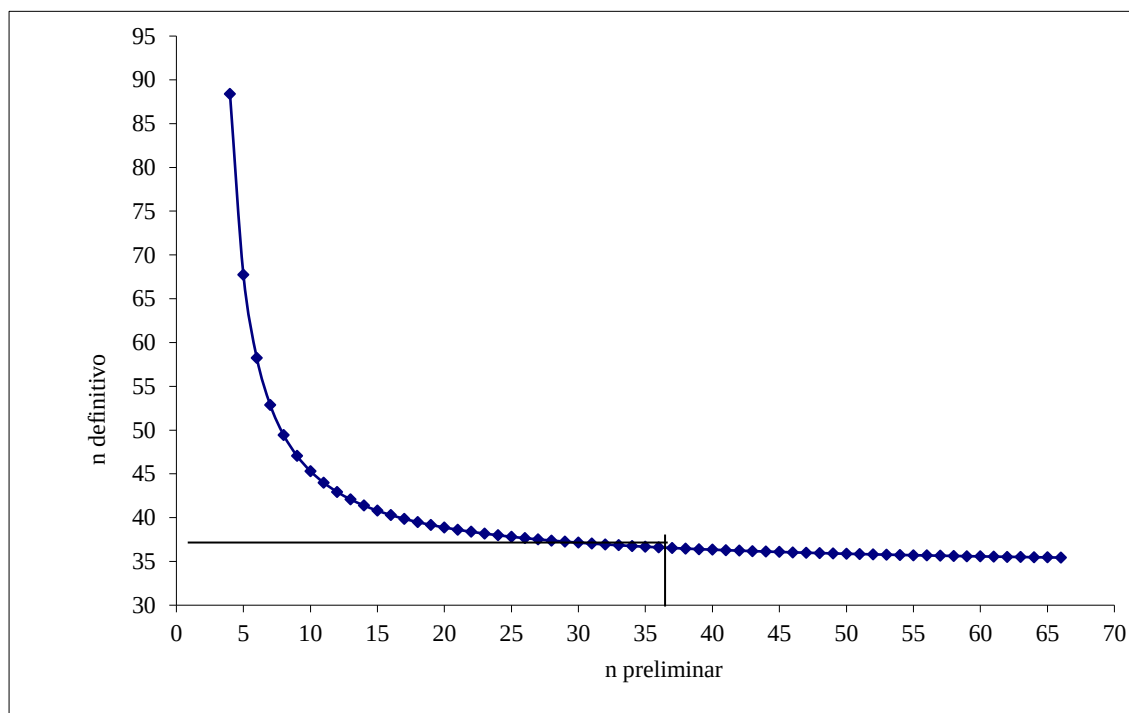
- El valor de N = 3000 parcelas, se obtiene de:

$$N = \frac{A}{a} = \frac{\text{area total}}{\text{area parcela}} = \frac{300 \text{ ha} \times 10,000 \text{ m}^2/\text{ha}}{1,000 \text{ m}^2 / \text{parcela}} = 3,000 \text{ parcelas}$$

- Error permisible = 20%
- Nivel de confianza = 95%
- Como utilizamos 5 parcelas como muestra inicial por estrato, tenemos $gl = 5 - 1 = 4$, con este dato tenemos que el valor de $t = 2.776$.
- Con esta información obtenemos el primer dato de número de parcelas:

$$n = \frac{2.776^2 \times (59.07)^2}{20^2 + \frac{2.776^2 \times (59.07)^2}{3,000}} = 67.7 \approx 68$$

- Luego se continúa probando con diferentes grados de libertad, en el rango comprendido entre 4 y 67.



Analizando el anterior gráfico, se observa que el tamaño de muestra definitivo es de 36 parcelas. Otra forma de determinar el tamaño de muestra definitivo es evaluando la ecuación de tamaño de muestra definitiva con $gl = 36$ (promedio entre 4, gl inicial, y 67, gl del primer cálculo de tamaño de muestra):

$$n = \frac{2.03^2 \times (59.07)^2}{20^2 + \frac{2.03^2 \times (59.07)^2}{3,000}} = 36.53 \approx 37$$

Para la distribución de las parcelas se utilizará la distribución proporcional, ya que se tomará en cuenta el área de cada estrato. La distribución se realizará de la siguiente forma:

Estrato A:	n_1	$= n (P_1) = 36 (0.27)$	$= 10$
Estrato B:	n_2	$= n (P_2) = 36 (0.33)$	$= 12$
Estrato C:	n_3	$= n (P_3) = 36 (0.40)$	$= 14$

2.2.5 Análisis estadístico de los resultados

Ejemplo:

Una población de 2,500 trozas ubicadas en el patio de una industria forestal que trabaja con diámetros pequeños fue dividida en tres estratos, según el lugar de origen: San Pablo, San Miguel y Candelaria. Para estimar el volumen total en el patio se utilizó un muestreo estratificado tomando una muestra de 55 trozas. La muestra se distribuyó proporcionalmente al número de trozas en cada estrato. En cada estrato se aplicó un muestreo aleatorio, cuyos resultados aparecen en el cuadro siguiente.

Origen	Número de trozas (N _i)	Proporción (P _i)	Tamaño de muestra (n _i)	Volumen promedio m ³ /troza (V _i)	Desviación estándar m ³ /troza (s _i)	Error estándar m ³ /troza (S _{$\bar{x}$ i})	Volumen total m ³ (τ _i)
San Pablo	1350	0.54	30	0.251	0.1042	0.0188	338.85
San Miguel	700	0.28	15	0.164	0.098	0.0250	114.80
Candelaria	450	0.18	10	0.114	0.055	0.0172	49.50
	N = 2500						503.15

Con base en los anteriores resultados, calcule el volumen promedio por troza, el volumen total en el patio y sus límites de 95% de confianza.

- Volumen promedio por troza:

$$\bar{x}_p = \frac{1}{N} \sum_{i=1}^k \bar{x}_i \times N_i = \frac{1}{2500} [(0.251 \times 1350) + (0.164 \times 700) + (0.11 \times 450)] = 0.2013 \text{ m}^3/\text{troza}$$

- Error estándar del volumen promedio por troza:

$$S_{\bar{x}_p} = \sqrt{\sum_{i=1}^k P_i^2 \times S_{\bar{x}_i}^2} = \sqrt{(0.54^2 \times 0.0188^2) + (0.28^2 \times 0.025^2) + (0.18^2 \times 0.0172^2)} = 0.01271 \text{ m}^3/\text{troza}$$

- Error de muestreo al 95% de confianza:

$$E = S_{\bar{x}_p} \times t_{(\alpha/2, n-k)} = E = 0.01271 \times 2.01 = 0.026 \text{ m}^3 / \text{troza}$$

n = tamaño de la muestra

k = número de estratos

$$E\% = \frac{E}{\bar{x}_p} \times 100 = E\% = \frac{0.026}{0.2013} \times 100 = 12.7\%$$

- Volumen total en el patio:

$$\hat{\tau} = \bar{x}_p \times N = 0.2013 \times 2500 = 503.25 \text{ m}^3$$

- Intervalo de (1-α)% de confianza:

$$IC_{(\hat{\tau})} = N \times \left[\bar{x}_p \pm S_{\bar{x}_p} \times t_{\alpha/2, n-k} \right] = 2,500 \times [0.2013 \pm 0.01271 \times 2.01]$$

$$439.38 \text{ m}^3 \leq \tau \leq 567.12 \text{ m}^3$$

2.3 MUESTREO SISTEMÁTICO

2.3.1 Definición

Scheaffer et. Al (1986) consideran que una muestra obtenida al seleccionar aleatoriamente un elemento de los primeros k elementos en el marco y después cada k -ésimo elemento, se denomina muestra sistemática de 1 en k .

Orozco & Brumér (2002) indican que el muestreo sistemático es el más utilizado en inventarios forestales en bosques naturales tropicales, a pesar de que en teoría este tipo de muestreo no está basado en las leyes de la probabilidad, lo que no permite calcular con precisión el error de muestreo. Sin embargo, en la práctica casi todos los inventarios realizados con muestreo sistemático son analizados utilizando las ecuaciones de muestreo simple aleatorio, debido a que es posible introducir procedimientos especiales que validen el cálculo del error muestreo, por ejemplo, la selección al azar de la primera unidad muestral, parcela o transecto.

De acuerdo con Batista (2002), al localizar las unidades muestrales sistemáticamente en el campo, se consigue una mejor cobertura espacial, en comparación con el muestreo simple aleatorio, lo que es interpretado como una mejor representatividad espacial. Malleux (1982) coincide en señalar que el muestreo sistemático generalmente da resultados más precisos que el muestreo simple al azar, debido a que el área es cubierta en una forma más regular y proporcional; además, los cálculos generan un aproximado del máximo error de muestreo, y no del promedio del error de muestreo como es el caso del muestreo al azar.

2.3.2 Ventajas

Las principales ventajas de los inventarios forestales con muestreo sistemático se resumen en:

- ❑ Rapidez y facilidad en la selección de los elementos de la muestra en la población, así como bajo costo por la simplicidad de la selección. Los inventarios son usualmente más rápidos y baratos, ya que se reduce el tiempo invertido en la localización de las parcelas de medición y en el desplazamiento de una parcela a otra.
- ❑ Mejor organización y control en el trabajo de campo. El área puede ser mapeada en el terreno, dado que las cuadrillas de campo deben recorrer toda el área a intervalos fijos y con un rumbo definido; esto permite identificar estratos de bosque, áreas de protección y el área efectiva de manejo. Con este sistema se puede recopilar la información necesaria para preparar mapas de curvas a nivel utilizando técnicas de interpolación espacial.
- ❑ Ninguna sucesión grande de elementos en la población queda sin representación. El muestreo sistemático permite hacer estimaciones confiables de medias poblacionales y de totales (por ejemplo, volumen total), ya que la muestra se distribuye en toda la población. En muchos casos, las estimaciones son mejores que las obtenidas con un muestreo simple aleatorio.

2.3.3 Desventajas

Las principales desventajas son:

- ❑ Desde el punto de vista teórico, no proporciona un error de muestreo exacto, en comparación cuando los parámetros se examinan como muestras seleccionadas al azar.

- ❑ La estimación del error estándar es más compleja y sólo es una aproximación.
- ❑ El error estándar depende del orden o distribución de las parcelas en el terreno.
- ❑ Si el muestreo coincide con alguna distribución periódica de la población, la estimación de la media población resulta ser muy mala.
- ❑ Las estimaciones del total y la media aritmética tienen un pequeño sesgo (conforme aumenta el tamaño de la muestra, las estimaciones tienden a alejarse de los valores poblacionales). Sin embargo, este sesgo se puede reducir si se utiliza un muestreo sistemático con arranque aleatorio.

2.3.5 Procedimiento para realizar un muestreo sistemático con arranque aleatorio

1. Seleccione un intervalo de muestro k , la mejor alternativa es seleccionar k igual al tamaño de la población (N) dividido entre el tamaño de la muestra (n). Esto es, $k = N / n$.
2. Seleccione aleatoriamente un número entero i (arranque o inicio aleatorio) comprendido entre 1 y k ($1 \leq i \leq k$)
3. Mida primero la parcela o unidad de muestreo seleccionada aleatoriamente (i) en paso 2, luego de manera rígida o sistemática tome la unidad $i + k$, que está k lugares del i -ésimo en la lista, el $i + 2k$ que está $2k$ después y así sucesivamente hasta completar el tamaño n de la muestra. Note que la tabla de número aleatorios u otro mecanismo de selección se emplea una sola vez, en i .

Ejemplo:

En un bosque con área efectiva (A) igual a 20 ha se desea estimar el número de árboles por hectárea (N_a). Para hacer la estimación se utilizará el muestreo sistemático con parcelas de 0.5 ha con una intensidad del 10%. El tamaño de la población (N) es, entonces, 40 y el tamaño de la muestra (n) es 4. El proceso de selección de las parcelas de inventario utilizando un arranque aleatorio es el siguiente:

1. Calcule el intervalo de muestreo k , donde $k = N / n = 40 / 4 = 10$.
2. Seleccione aleatoriamente un número entero entre 1 y 10 $\rightarrow 3$
3. Las parcelas que forman la muestra son, entonces: 3, $3 + 10$, $3 + (2)(10)$ y $3 + (3)(10)$, o sea: 3, 13, 23 y 33.

En la siguiente figura, se presenta el plano del bosque de 20 ha dividido en parcelas de 0.5 ha. Los números representan el número asignado a cada parcela para proceder a realizar el muestreo sistemático con una intensidad del 10%.

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40

Cálculos:

Si en el inventario forestal de un bosque con área efectiva (A) se utilizan unidades de muestreo de tamaño “a” en hectáreas, se tiene que el tamaño de la población $N = A/a$. Si el tamaño de la muestra es “n”, la fracción de muestreo (f) es igual a n/N .

Si $x_1, x_2, x_3, \dots, x_n$ son los resultados de la evaluación de “n” unidades de muestreo seleccionadas sistemáticamente, al dividir cada valor $x_1, x_2, x_3, \dots, x_n$ entre “a” se obtienen los valores extrapolados a hectárea. Una vez extrapolados los datos por parcela a hectárea se obtiene:

- El promedio por hectárea ($\bar{x}$):
$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$
- El total estimado ($\hat{\tau}$) en el área efectiva del bosque (A): $\hat{\tau} = \bar{x} \times A$

La estimación del error estándar es problemática en un muestreo sistemático dado que no hay un estimador insesgado; solo se pueden calcular aproximaciones del verdadero valor. Orozco & Brumér (2002) citan dos aproximaciones:

1. Si el técnico forestal está convencido de que la selección sistemática dio como resultado una selección de parcelas semejante a un muestreo aleatorio, se puede asumir que el error estándar es similar al que se obtiene con un muestreo simple aleatorio. Por lo tanto el error estándar ($S_{\bar{x}}$) se calcula con las ecuaciones del muestreo simple aleatorio.
2. Si el profesional forestal sospecha que las parcelas no siguen una distribución aleatoria sobre el terreno, el error estándar se calcula utilizando una aproximación de las diferencias al cuadrado de observaciones sucesivas. Con este método, primero se calcula la suma de las diferencias al cuadrado de observaciones sucesivas:

$$\sum_{i=1}^n (x_{i+1} - x_i)^2$$

En esta expresión, a la observación de la segunda parcela en la secuencia de selección, se le resta el valor de la primera y el resultado se eleva al cuadrado. Al valor de la tercera parcela seleccionada se le resta el valor de la segunda, el resultado se eleva al cuadrado y se le suma al resultado de la diferencia al cuadrado anterior, y así sucesivamente hasta llegar a la observación de la penúltima parcela.

Una vez calculada la suma de las diferencias al cuadrado de observaciones sucesivas, se calcula una desviación estándar (S_{ps}) aproximada con la ecuación:

$$S_{ps} = \sqrt{\frac{\sum_{i=1}^n (x_{i+1} - x_i)^2}{2(n-1)}}$$

el error estándar de la media es igual a:

$$S_{\bar{x}} = \frac{S_{ps}}{\sqrt{n}} \quad \text{o} \quad S_{\bar{x}} = \frac{S_{ps}}{\sqrt{n}} \times \sqrt{1 - \frac{n}{N}}$$

El error de muestreo absoluto (E) y el relativo (E%) para la estimación de la media poblacional con una confiabilidad de $(1-\alpha)100\%$ se calcula con $n-1$ grados de libertad y es igual a:

$$E = S_{\bar{x}} \times t_{\alpha/2, n-1} \quad E\% = \frac{E}{\bar{x}} \times 100$$

Los límites de confianza para la estimación de la media poblacional por unidad de área con una confiabilidad de $(1-\alpha)100\%$ están dados por:

$$IC_{(\mu)} = \bar{x} \pm S_{\bar{x}} \times t_{\alpha/2, n-1}$$

Los límites de confianza para la estimación del total ($\hat{\tau}$) en la población con una confiabilidad de $(1-\alpha)100\%$ están dados por:

$$IC_{(\tau)} = A \times \left[\bar{x} \pm S_{\bar{x}} \times t_{\alpha/2, n-1} \right]$$

Al igual que con el muestreo simple aleatorio, si este diseño de muestreo (sistemático) se utiliza en inventarios cuyas unidades de muestreo no son parcelas, por ejemplo, en inventarios cuyas unidades de muestreo son árboles, plántulas en un vivero, semillas, algún producto de un árbol o una especie animal, el total de la población se calcula multiplicando el promedio por unidad de muestreo, por el tamaño de la población (N). En este caso, las ecuaciones del total y sus límites de confianza se transforman en:

$$\hat{\tau} = \bar{x} \times N$$

$$IC_{(\tau)} = N \times \left[\bar{x} \pm S_{\bar{x}} \times t_{\alpha/2, n-1} \right]$$

Demostración:

- Parámetro de interés: $\tau = \mu \times N$
- Estimador: $\hat{\tau} = \bar{x} \times N$
- Varianza del estimador: $Var(\hat{\tau}) = s_{(\hat{\tau})}^2 = N^2 \frac{s^2}{n} \left[\frac{N-n}{N} \right],$

$$\text{como } Var(\hat{\mu}) = s_{(\bar{x})}^2 = \frac{s^2}{n} \left[\frac{N-n}{N} \right], \text{ entonces: } Var(\hat{\tau}) = s_{(\hat{\tau})}^2 = N^2 s_{(\bar{x})}^2$$

- Error estándar del estimador: $s_{(\hat{\tau})}^2 = \sqrt{s_{(\hat{\tau})}^2} = N \times s_{\bar{x}}$
- $IC_{(\tau)} = \tau \pm s_{\hat{\tau}} \times t_{\alpha/2, n-1}$, como $\hat{\tau} = \bar{x} \times N$ y $s_{(\hat{\tau})} = N \times s_{\bar{x}}$, la ecuación de intervalo de confianza para el total queda de la siguiente forma:

$$IC_{(\tau)} = (N \times \bar{x}) \pm N \times s_{\bar{x}} \times t_{\alpha/2, n-1} = IC_{(\tau)} = N \left[\bar{x} \pm s_{\bar{x}} \times t_{\alpha/2, n-1} \right]$$

Ejercicio:

En un inventario forestal de un bosque natural de 500 ha, se calculó que el área efectiva de manejo era de 370 ha, mientras que las restantes 130 ha fueron clasificadas como zonas de protección por pendiente (mayor a 70%), y zonas de protección de ríos y fuentes de agua. Para planificar el manejo del bosque, se realizó un inventario forestal general.

Para realizar este inventario se utilizaron parcelas rectangulares de 30m × 100m (0.3 ha). Se seleccionaron sistemáticamente 11 parcelas, los resultados de volumen comercial (Vc) por parcela para árboles con diámetro superior a 60 cm se presentan en el cuadro siguiente:

Parcela	1	2	3	4	5	6	7	8	9	10	11
Vc (m ³ /parcela)	3.6	3.15	4.65	4.44	5.1	3.81	4.14	2.7	3.54	4.89	4.23
Vc (m ³ /hectárea)	12.0	10.5	15.5	14.8	17.0	12.7	13.8	9.0	11.8	16.3	14.1

Con base en estos resultados, calcule el volumen comercial total y el error de muestreo con un nivel de 95% de confianza, y los límites de confianza del total con igual confiabilidad. Para calcular el error estándar, utilice el método de las diferencias al cuadrado de observaciones sucesivas.

Solución:

a) Datos básicos

Área efectiva(A)	=	370 ha
Tamaño de parcela (a)	=	0.3 ha
Tamaño de la población (N)	=	1233
Tamaño de muestra (n)	=	11
Fracción de muestreo (f)	=	0.0089
t _(0.05/2,10)	=	2.228

b) Cálculos

$$\sum_{i=1}^n x_i = 147.5 \text{ m}^3$$

$$\sum_{i=1}^n (x_{i+1} - x_i)^2 = (10.5 - 12.0)^2 + (15.5 - 10.5)^2 + \dots + (14.1 - 16.3)^2 = 108.25 \text{ u}^2$$

$$\text{Volumen promedio por hectárea: } V_c = \frac{147.5 \text{ m}^3}{11 \text{ ha}} = 13.401 \text{ m}^3 / \text{ha}$$

Desviación estándar aproximada del volumen comercial por hectárea:

$$S_{ps} = \sqrt{\frac{108.25}{2(11-1)}} = 2.326 \text{ m}^3 / \text{ha}$$

Error estándar de la media del volumen comercial por hectárea:

$$S_x = \frac{2.326}{\sqrt{11}} = 0.701 \text{ m}^3 / \text{ha}$$

Error de muestreo absoluto y relativo al 95% de confianza:

$$E = S_{\bar{x}} \times t_{\alpha/2, n-1} = 0.701 \times 2.228 = 1.563 \text{ m}^3$$

$$E\% = \frac{1.563}{13.401} \times 100 = 11.66\%$$

Total de volumen comercial ($\hat{\tau}$) en el área de bosque efectiva:

$$\hat{\tau} = A \times V_c = 370 \times 13.401 = 4958.37 \text{ m}^3$$

Límites de 95% de confianza para la estimación del volumen total:

$$IC_{(\tau)} = A \times \left[\bar{x} \pm S_{\bar{x}} \times t_{\alpha/2, n-1} \right] = 370 \times [13.401 \pm 0.701 \times 2.228]$$

$$4382.97 \text{ m}^3 \leq \tau \leq 5533.77 \text{ m}^3$$

CAPITULO III

EJEMPLOS DE APLICACIÓN

Plan de muestreo de huevos de chinche salivosa (*Aeneolamia* spp.). **Badilla, F.; Hidalgo, H.; Salguero, V. (1999)**

Objetivo:

El muestreo de huevos de chinche salivosa permite detectar áreas de alta infestación y predecir las poblaciones de adultos para el siguiente período lluvioso. Con esta información se puede programar las medidas preventivas de control y orientarlas a las áreas de alta infestación. El muestreo de huevos también es necesario en actividades de investigación.

Época de muestreo:

El muestreo de huevos de chinche salivosa se realiza tomando muestras de suelo, las cuales son sometidas al proceso de extracción de los huevos. Este muestreo debe realizarse inmediatamente después del corte, debido a que en ese momento existe una mayor facilidad para la toma de muestras por la ausencia de caña parada. Esto permite una distribución adecuada de las muestras. Esto, cuando el muestreo se hace para detectar áreas de infestadas durante la época seca.

¿Dónde muestrear?

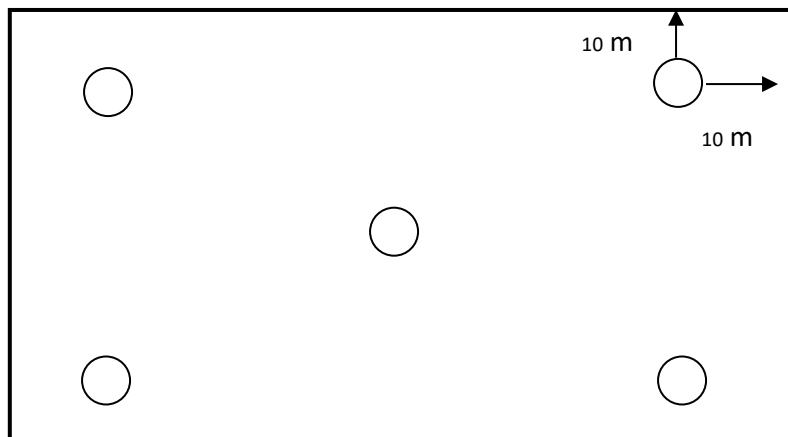
Debe muestrearse los pantes afectados por la plaga el año anterior y los pantes próximos que reúnan condiciones adecuadas para el desarrollo de la misma, por ejemplo: pantes con problemas de drenaje y suelos pesados. Las muestras de suelo se toman en el centro de la macolla, es decir, entre las cañas.

Número de muestras

El número de muestras a tomar es de 5 en un pante con una extensión máxima de 5 ha. En pantes mayores a 5 ha. se toma 1 muestra adicional por cada hectárea.

Distribución de las muestras

Las muestras deberán distribuirse sistemáticamente, 1 en cada esquina del pante y la última en el centro. Las 4 muestras de las esquinas deberán estar a 10 metros de la orilla.



Toma de la muestra

Para tomar las muestras de suelo se usa un marco de hierro de 30×30 cm y 6 cm de altura o bien un cilindro de hierro de 8 cm de diámetro. Cuando se usa el marco, debe profundizarse en el centro de una macolla, las raíces y residuos de tallos de la macolla se sacan con una pala de jardinero, tratando que se quede la menor cantidad de suelo posible en ellos. Con la misma pala se colecta el suelo que se encuentre hasta la profundidad de 6 cm y se coloca en una bolsa plástica. Si se usa el cilindro, debe graduarse a 4 cm de profundidad teniendo el cuidado de ubicarlo entre la macholla en lugares sin rastros.

Identificación de la muestra

Todas las muestras deben llevar la siguiente identificación: finca, fecha del muestreo, lote, pante, localidad, colector.

Homogenización y pesado de las muestras

Cuando se utiliza el marco de hierro, antes de realizar el análisis de laboratorio debe de homogeneizarse las 5 muestras. Esta muestra homogeneizada se pesa y se obtiene el peso promedio por muestra. Luego se extrae una submuestra de 250 gramos, que será lo único que se analizará en el laboratorio.

Si se usa cilindro, no es necesario homogeneizar ni tomar una submuestra, sino que se procesa el suelo de las 5 muestras.

3.2 Cálculo del tamaño de la muestra para muestreo de chinche hedionda

En un muestreo preliminar realizado para cuantificar el número de ninfas + adultos de Chinche Hedionda se recolectaron los siguientes datos:

Unidad de muestreo	1	2	3	4	5	6	7	8	9	10
Individuos	22	5	19	49	36	17	6	4	12	15

Tamaño de la muestra piloto (n_1) = 10

Media = 18.50

Desviación estándar = 14.35

Varianza = 206.10

Tamaño de muestra definitivo

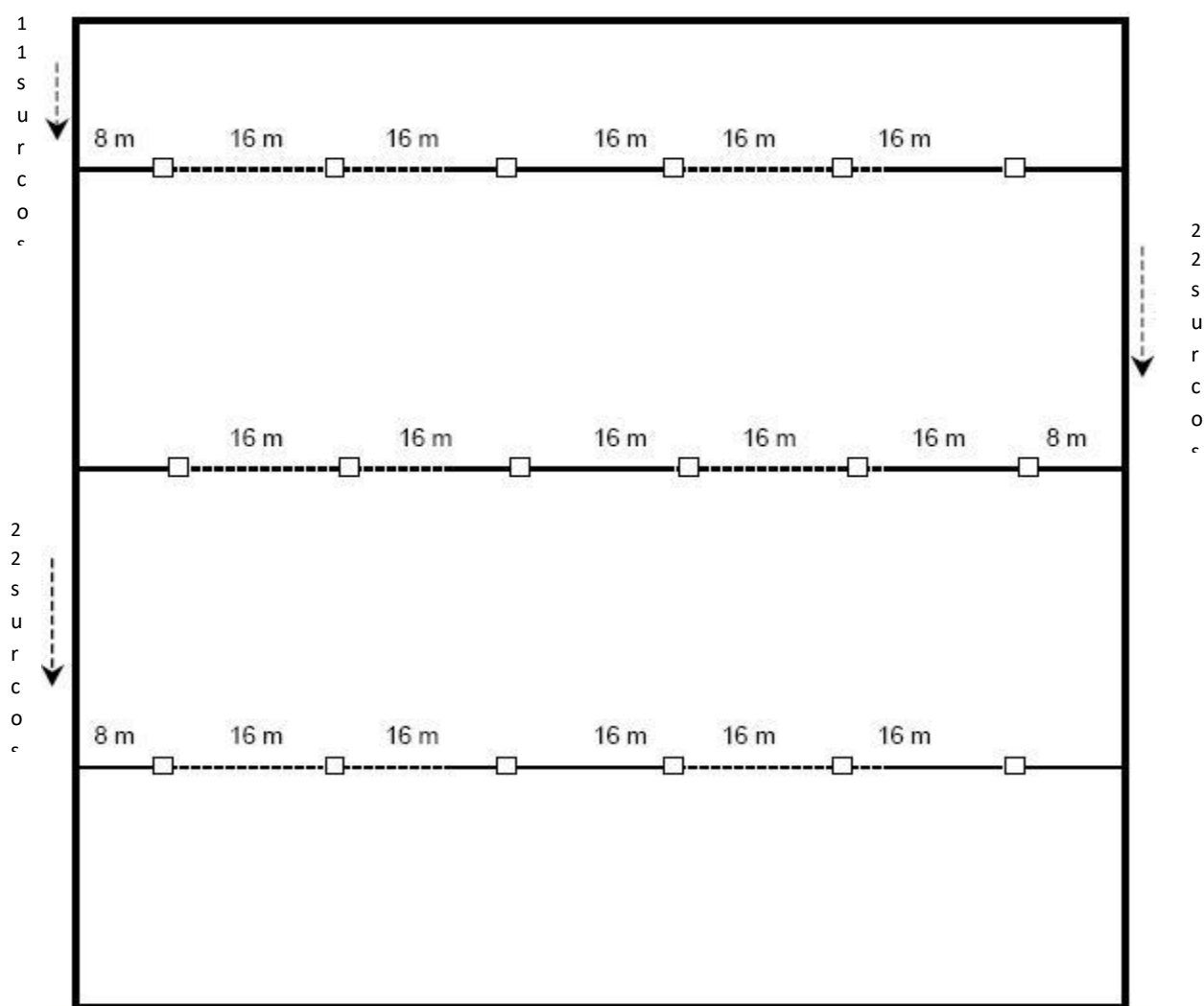
$$n = \left(\frac{14.35}{0.20 * 18.50} \right)^2 = 15$$

3.3 Plan de muestreo para la cigarrilla de las raíces (*Mahanarva fimbriolata*, Stal. 1854) de caña de azúcar.

Stingel, E.; Bento, J. (2005)
Piracicaba, Estado de São Paulo.

Estos autores llegaron a la conclusión que el plan de muestreo más adecuado para la evaluación de las poblaciones de cigarrilla de las raíces (una especie de chinche salivosa) en el cultivo de la caña de azúcar, consiste en la toma de 18 muestras de 1 metro lineal por hectárea, distribuidas conforme un esquema E de caminamiento. Para que la muestra sea representativa, es fundamental estandarizar la toma de la muestra, 1 metro lineal sobre el surco y 0.5 m de cada lado del surco a muestrear.

En la siguiente figura se muestra la forma como recomiendan distribuir las unidades de muestreo en el campo.



□ Punto de muestreo

3.4 Estimación de la suficiencia muestral para evaluar intensidad de infestación de *Diatraea* spp. en caña de azúcar.

Aleixo, J.A.; Lira, A.; Souza, A. F.; Carrano, A. F.; Marques, E. J. (1997)

Metodología

1. Lugar de la investigación: Ingenio Trapiche, municipio de Serinhaém, Pernambuco.
2. Variedad de caña: CB 45-3, caña en pie.
3. Método de muestreo: Simple aleatorio (por tratarse de una única variedad)
4. Fecha de realización del estudio: agosto 1991.
5. Tamaños de muestra evaluados (por hectárea)

Tamaño de muestra (número de tallos)	Número de unidades muestrales tomadas (parcelas)	Total de tallos por hectárea
10	20	200
20	17	340
40	15	600
60	12	720
80	10	800
100	8	800

Observaciones:

- En cada hectárea fue dejado un borde de 10 metros, usándose únicamente el área central del cañal, para la toma de las muestras.
 - En la selección del número de unidades muestrales (parcelas) para cada tamaño de muestra, se consideró el principio de que muestras de menor tamaño deber ser muestreadas en mayor número de veces, y viceversa.
6. Cada conjunto de tallos (unidad muestral) fue amarrado en maletas y etiquetado con el número de unidad muestral.
 7. Cada tallo fue cortado en la base y en la punta, y transportado sin hojas. Las maletas fueron analizadas individualmente.
 8. Fueron cronometrados el tiempo de operación de colecta de parcelas en el campo (T_1) y el tiempo de observación, abertura longitudinal y evaluación de infestación por caña (T_2). El tiempo total TT fue obtenido por la suma de T_1 y T_2 .
 9. La evaluación de la intensidad de infestación (II%) fue realizado según **Macedo y Botelho** (1998), cuya ecuación está expresada por:

$$II\% = \left(\frac{NEB}{NTE} \right) \times 100,$$

siendo:

NEB = Número de entrenudos con daño de broca
NTE = Número total de entrenudos.

10. El proceso de selección del tamaño ideal de la parcela fue el de la EFICIENCIA RELATIVA (ER%). (Covas y Christensen, 1945; Freese, 1962; Machado y Albertini, 1973, Scolforo et al, 1993). La ER% de cada parcela fue comparada con relación a las demás, a través de la ecuación:

$$ER\% = \left(\frac{TT_i \times NPR_i \times EAR_i}{TT_j \times NPR_j \times EAR_j} \right) \times 100,$$

siendo:

TT_i = tiempo total medio de operación de colecta de la caña en parcelas del tipo i,
 TT_j = tiempo total medio de operación de colecta de la caña en parcelas del tipo j,
 Ambos tiempos expresados en horas.

NPR_i y NPR_j corresponden a los números de parcelas representativos de la población utilizando parcelas del tipo (i,j) definidos por:

$$NPR_x = \left(\frac{t^2_{(0.025)} \times CV\%^2}{EA\%^2} \right), \quad x = i, j$$

siendo:

$t_{(0.025)}$ = valor en la distribución de t de Student a 5% de probabilidad. Los grados de libertad están dados por: número de parcelas -1.
 $CV\%$ = coeficiente de variación expresado en porcentaje,
 $EA\%$ = error de muestreo adoptado (10%),

EAR_i y EAR_j son los errores muestrales reales (ocurridos en el campo), usando parcelas del tipo (i,j), definidos por:

$$EAR_x = \frac{t_{(0.025)} \times S_{\bar{x}}}{II\%}, \quad x = i, j; \quad S_{\bar{x}} = \frac{s}{\sqrt{n}}$$

siendo:

$S_{\bar{x}}$ = error estándar de la media,
 $II\%$ = intensidad media de infestación
 s = desviación estándar
 n = número de unidades muestrales tomadas (parcelas)

11. Interpretación de los resultados.

- Cuando el valor de ER% sea menor que 100, se concluye que el tipo de parcela usado en el numerador es más eficiente que el tipo usado en el denominador, y viceversa. Cuando sea igual o próximo a 100, los tipos de parcelas son considerados semejantes.
- Generalmente se considera un tipo de parcela como estándar (100% eficiente), y se hace la comparación con los otros. En este trabajo no se utilizó una parcela estándar, siendo todos los tipos de parcelas comparados entre sí.
- Una gran ventaja de la prueba de la ER% es que además de considerar las medidas de variación también considera costos, buscando encontrar un tamaño ideal de parcela con buena precisión y bajo costo de toma de los datos.

Resultados obtenidos

Los valores utilizados en los cálculos de ER% de los diferentes tipos de parcela (número de tallos), se encuentran en la Tabla 1. El tipo de parcela compuesto por 20 tallos fue el más eficiente entre los usados en este trabajo (Tabla 2). Para comparar la ER% de este tipo de parcela con cualquier otro, se debe considerar la columna que tiene de encabezado el número 20. Se concluye que ese tipo de parcela fue 190.61 % más eficiente que el tipo de parcela donde fueron usados 10 tallos, y así sucesivamente. Ahora, considerando la línea que inicia con el número 20, se observa que el tipo de parcela que usó 10 tallos sólo es 34.37% eficiente cuando se compara con el tipo que usó 20 tallos. Cualquier comparación puede realizarse a través de la Tabla 2.

Tabla 1 Resultados muestrales de las variables utilizadas en el cálculo de la eficiencia relativa, en los diversos tamaños de parcelas estudiadas.

TP	NP	II%	TT	CV%	EAR%	NPR
10	20	4.194	7.28	69.70	32.59	213
20	17	6.872	12.29	39.90	20.19	70
40	15	7.847	24.25	37.10	20.52	64
60	12	6.291	33.00	26.90	17.08	36
80	10	6.428	42.25	33.00	23.56	56
100	8	4.368	58.51	37.40	31.19	80

Tabla 2 Resultados de las eficiencias relativas de cada tamaño de parcelas (número de tallos por hectárea) comparado con los demás.

Tamaño de la parcela (numerador)	Tamaño de la parcela (denominador)					
	10	20	40	60	80	100
10	100.00	290.61	158.77	249.06	90.67	35.50
20	34.37	100.00	54.58	85.62	31.17	11.61
40	62.98	183.22	100.00	156.87	57.11	22.36
60	40.15	116.80	63.75	100.00	36.40	14.25
80	110.29	320.82	175.10	274.73	100.00	39.15
100	281.69	861.67	447.23	701.75	255.42	100.00

Referencias:

TP	=	tamaño de la parcela (número de tallos),
NP	=	número de parcelas medidas,
II%	=	intensidad de infestación en porcentaje,
TT	=	tiempo total medio de operación de la toma de la muestra (horas),
CV%	=	coeficiente de variación en porcentaje,
EAR%	=	error muestral real,
NPR	=	número de parcelas representativo de la población.

Basándose en los resultados obtenidos en este trabajo, se concluye que para la variedad de caña CB 45-3, en caña en pie, en las condiciones muestreadas, el mejor tamaño de muestra es el de 20 tallos.

En función de este resultado, se obtuvo un límite de confianza en cuanto a la intensidad de infestación igual a $6.872\% \pm 1.388\%$. Por tanto, para los datos estudiados la media de la intensidad de infestación por hectárea estuvo en torno de 6.872%, pudiendo variar entre 5.484% y 8.260%, con un 95% de confianza y con un error de muestreo real de 20.19%. Para reducir ese error a aproximadamente 10%, deberían ser tomadas 53 parcelas más.

3.5 Norma técnica obligatoria nicaragüense para el muestreo de productos vegetales

* Versión html del archivo <http://www.mific.gob.ni/dtnm/conpub/2.pdf>

La Norma Técnica Nicaragüense 17 002 -02 ha sido preparada por el Comité Técnico de preparación y presentación de normas. Fue aprobada por la Comisión Nacional de Normalización Técnica y Calidad en sesión efectuada el día XXX de XXX del 2000.

1. OBJETO: Esta norma tiene por objeto establecer los procedimientos a seguir para la toma y preparación de las muestras en productos vegetales para realizar los análisis fitosanitarios con fines de certificación.

2. CAMPO DE APLICACIÓN: La presente norma es aplicable a todas las instituciones, personas naturales y jurídicas dedicadas a la exportación, distribución de productos vegetales, de consumo humano.

3. DEFINICIONES

3.1 Equipo de muestreo. Son los diferentes instrumentos que se utilizan para extraer una cantidad de productos de un empaque o envase.

3.2 Equipo de laboratorio. Se refiere a los diferentes instrumentos que Se utilizan para realizar el análisis de diferentes sustancias.

3.3 Establecimiento. Lugar autorizado por la Autoridad de aplicación, donde sé, comercializa y almacena los vegetales.

3.4 Funcionario encargado del muestreo. Persona capacitada en materia de procedimientos de muestreo y autorizada por las autoridades competentes para tomar muestras.

3.5 Lote. Es una cantidad de productos de la misma naturaleza y procedencia, que tengan características presumiblemente uniformes.

3.6 Muestra. Es la cantidad de material cuya composición debe representar estadísticamente la totalidad del material de donde se tomó, con el fin de ser analizada en el laboratorio.

3.7 Muestreo. Es el conjunto de operaciones que se lleva a cabo con el objeto de extraer muestras representativas de un determinado lote.

3.8 Muestra analítica. Es la porción de producto que ha de analizarse a partir de la muestra de laboratorio.

3.9 Muestra simple ó elemental. Es una determinada cantidad de material que se extrae de un sublote o lote.

3.10 Muestra compuesta o global. Es aquella muestra obtenida por homogeneización de diferentes muestras simples o elementales, con el fin de garantizar una muestra representativa de la totalidad del material.

3.11 Muestra final. Es la parte representativa obtenida de la división en triplicado de la muestra compuesta.

3.12 Muestra de laboratorio. Es una parte de la muestra final que se envía al laboratorio para el análisis, que puede utilizarse como un todo o sub-dividirse en porciones representativas, si así lo exige la legislación nacional.

3.13 Sublote. Es cada una de las partes aproximadamente iguales en que se divide un lote en forma imaginaria o real para tomar de cada una de ellas una o varias muestras iguales.

3.14 Tamaño de la muestra. Número de unidades, o cantidad de material que constituye la muestra.

3.15 Vegetales. Son productos que se consumen en estado natural o con un mínimo de procesamiento (crudas). Ejemplo: Zanahoria, rábano, fresas, tomates, apio, coliflor y repollo.

4. MATERIALES, EQUIPOS Y REACTIVOS a) Etiquetas b) Marcadores c) Libretas d) Cuchillas y/o tijeras para podas e) Actas de muestreo f) Guantes g) Hieleras h) Lupas 10x i) Alcohol 70% j) Viales k) Lámparas de mano.

5. TOMA DE MUESTRAS.

5.1 Métodos de muestreo. El muestreo para certificación de los productos vegetales, deberá ser ejecutado por personal autorizado, quienes deberán de estar provistos de todo el equipo y materiales necesarios para dicha actividad, debiendo solicitar la colaboración de la empresa propietaria del producto a muestrear, y deberán de seguirse las siguientes instrucciones:

- Realizar el muestreo en presencia de un representante calificado de la empresa.
- Identificar el o los productos a muestrear en sus empaques originales (cajas, sacos). Estos serán debidamente sellados después del muestreo, si existen empaques dañados ó abiertos se hará un muestreo adicional, informando en el acto al representante de la empresa.
- Identificar los lotes de los productos vegetales a muestrear.
- Determinar el número y tamaño de las submuestras a tomar de cada lote.
- Seleccionar al azar las unidades de cada lote a muestrear.
- Se debe de muestrear la parte del producto que es comerciable.
- Deberán de utilizarse empaques resistentes de acuerdo al tipo de producto muestreado. Homogeneizar y dividir la muestra en tres partes.
- Las muestras serán debidamente identificadas con etiquetas o colillas en la que se detallará: Número de la muestra, nombre del vegetal, lugar en donde se tomo la muestra, cantidad de producto muestreado y nombre del propietario. Etiquetar cada muestra final, sellarla y distribuirla una a la empresa, otra para el laboratorio y una testigo.
- Etiquetar cada lote muestreado.
- Levantar el acta de muestreo y entregar una copia a la empresa.
- Llenar hoja de envío al laboratorio

5.2 Tamaño de la muestra. La selección del tamaño de la muestra depende del tipo, descripción y tamaño mínimo de la muestra a como se indica en la siguiente descripción:

Productos de origen vegetal, descripción y tamaño mínimo de las muestras

Clasificación de los productos	Ejemplos	Tamaño mínimo de la muestra
1. Productos frescos de tamaño pequeño unidades de 25 gr.	Apio, lechuga, espinacas, guisantes y aceitunas.	1 kg
2. Productos frescos de tamaño medio, unidades de 25 a 250 gr.	Tomate, manzanas, naranjas, peras, melocotones, otras.	1 Kg. (10 unidades, al menos).
3. Productos frescos de tamaño grande, unidades 250 gr.	Coles, pepinos, racimos de uvas, melones, piña, papaya y remolacha.	2 Kg. (5 unidades al menos)
4. Legumbres.	Soya, vainitas. Okra	1 Kg.
5. Hierbas aromáticas frescas.	Perejil, albahaca, otras.	0.5 Kg.
6. Especias.	Secas	0.1 Kg.
7. Leguminosas forrajeras y otros.	forrajes y piensos.	1 Kg.

6. PROCEDIMIENTO PARA LA TOMA DE MUESTRA

6.1 El procedimiento para la toma de muestra, tiene por objeto adquirir una muestra final representativa del lote, a fin de determinar la calidad y certificar todos los productos vegetales de consumo nacional, tanto como para la exportación e importación de los mismos. La muestra final se considera representativa del lote, cuando se ha obtenido según el procedimiento descrito a continuación:

6.1.1 Deberán evitarse la contaminación y el deterioro de las muestras en todas las fases, ya que podrían afectar los resultados de los análisis.

6.1.2 Deberán tomarse muestras por separado de cada lote.

6.1.3 Determinar la cantidad de muestras a recolectar.

6.1.4 Seleccionar el método de muestreo a emplear según el tipo de vegetal.

6.1.5 Cuando se recolectan muestras directamente en el campo de cultivo, no se deben tomar productos enfermos y las muestras deben tomarse durante el período de cosecha.

6.1.6 Se debe muestrear la parte del producto que es comerciable.

6.1.7 Se debe tener cuidado de no remover residuos superficiales en la muestra durante la recolección, el empaque o transporte de la muestra.

6.1.8 Se debe tomar y empacar la cantidad o el peso recomendado en el sitio del muestreo y no realizar submuestreos de lo empacado.

6.1.9 Los instrumentos que se utilizan para la toma de muestras deben de estar limpios y sin contaminación.

6.1.10 Se deben utilizar empaques y/o envases nuevos y en perfecto estado de limpieza.

6.1.11 Evitar la contaminación de las muestras causada por manos o ropas que hayan estado en contacto con plaguicidas.

6.1.12 Las muestras deben de transportarse adecuadamente y mantenerse así hasta que se realice el análisis.

6.2. Preparación de las porciones.

6.2.1 Fase Confirmativa. Preparar a partir de una muestra final, una muestra de laboratorio.

6.2.2 Fase de intervención. Preparar a partir de la muestra final tres muestras iguales de laboratorio (laboratory sample), para distribuir del siguiente modo.

- a) 1 (uno) ejemplar en poder del propietario del lote con una copia del acta, con la obligación de conservarla en perfecto estado para su utilización en prueba contradictoria si lo considera necesario.
- b) 1 (uno) ejemplar para el laboratorio para el análisis inicial y
- c) 1 (uno) ejemplar de referencia para la entidad que deba actuar en casos de discrepancia.

6.3 Acondicionamiento de las muestras finales.

6.3.1 Colocar cada muestra de laboratorio en un recipiente limpio de material inerte; que la proteja convenientemente ante cualquier factor de contaminación y daño que pueda derivarse del transporte.

6.3.2 Etiquetar y precintar después el recipiente de forma que sea imposible abrirlo o despegar la etiqueta sin deteriorar el precinto. Tomar así mismo todas las precauciones necesarias, para evitar cualquier

modificación de la composición de la muestra de laboratorio que pueda producirse durante su transporte o almacenamiento.

6.3.3 Las muestras deben de conservarse hasta la evaluación definitiva y hasta que se decida sobre el destino del lote. Si la calidad del producto es satisfactoria, los restos de las muestras tomadas deberán de eliminarse como residuos de cosecha.

6.3.4 Se les pondrán los sellos de las partes interesadas y se les colocará una tarjeta de identificación, todo ello en forma tal que no exista posibilidad de violación.

6.3.5 En las tarjetas deberán figurar los siguientes datos: a) Nombre del producto b) Lugar de procedencia del producto c) Peso o volumen del lote d) Lugar donde se toma la muestra e) Identificación de las muestras f) Fecha de la toma de muestra g) Nombre de la persona que tomó la muestra h) Nombre y domicilio de las partes interesadas i) Firma de las partes interesadas

6.4 Muestras elementales. En la medida de lo posible, recolectar las muestras en distintos puntos del lote. Consignar cualquier excepción de esta norma en la ficha o acta de toma de muestras. Los productos que estén total o fuertemente deteriorados no deberán ser objeto de toma de muestras.

6.5 Preparación de la muestra global. La muestra global se obtendrá mezclando las muestras elementales.

6.6 Preparación de la muestra final. La muestra global se podrá utilizar tal cual como muestra final. Si la muestra global fuere demasiado grande, la muestra final se podrá preparar a partir de ella por medio de un método de reducción adecuado. Dividir, por ejemplo, la muestra en cuatro partes, siguiendo los diagonales, eliminar dos cuartos opuestos, mezclar el resto; volver a dividir en cuatro y proseguir las operaciones hasta obtener la cantidad requerida.

7. IDENTIFICACION DE LA MUESTRA. Cada muestra de laboratorio deberá registrarse e identificarse correctamente y deberá ir acompañada con la etiqueta oficial y una ficha de muestreo, en la que se indique la naturaleza y origen de la muestra y la fecha y lugar de la toma de muestras, junto con toda la información complementaria que pueda ayudar al analista. Se asignará a la muestra de laboratorio un código exclusivo que se añadirá al registro de la muestra junto con todos los datos necesarios.

8. ACTA DE MUESTREO. El original del acta de muestreo se destinara a las oficinas de la autoridad competente, y una copia a la empresa propietaria del producto.

9. TRANSPORTE DE LA MUESTRA.

9.1 La muestra de laboratorio deberá colocarse en un envase limpio, e inerte, que ofrezca protección suficiente contra la contaminación exterior y contra los daños que puedan producirse en el traslado.

9.2 Etiquetar y precintar después el recipiente de forma que sea imposible abrirlo o despegar la etiqueta sin deteriorar el precinto.

9.3 La muestra, junto con la hoja de envío deberá ser entregada al laboratorio, de preferencia el mismo día de la toma. Si no es posible hacer esto, se deberán guardar de tal modo que no se altere el producto. El tiempo que dure almacenada no deberá ser mayor de 15 días.

9.4 Las otras dos porciones restantes, se almacenarán por un término de 30 días para efectos de discrepancia entre los interesados, en condiciones que no afecten el producto. Una vez vencido el período de almacenamiento, se deberán destruir las porciones mediante medios adecuados.

9.5 Fraccionamiento de la muestra. La muestra deberá dividirse en triplicado. Cada una de las fracciones se guardará, una para su ensayo inicial y las otras, en caso de resultar necesario, para verificar los resultados obtenidos en los análisis iniciales.

BIBLIOGRAFIA

Aleixo, J.A.; Lira, A.; Souza, A. F.; Carrano, A. F.; Marques, E. J. 1997. Estimativa da suficiência amostral para avaliar intensidade de infestação da *Diatraea* spp. em cana-de-açúcar. Pesquisa Agropecuária Brasileira. 32 (10): 1003-1007.

Badilla, F.; Hidalgo, H.; Salguero, V. Muestreo de huevos de chinche salivosa (*Aeneolamia* spp.) Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICANA). Noviembre de 1999.

Batista, J. L. F. 2002. Notas para acompanhar a disciplina: LCF 724 Métodos de amostragem em levantamentos florestais. Departamento de Ciencias Florestais, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP. Disponible en: <http://lmq.esalq.usp.br/syllabvs/lcf5764/notas.pdf>

Boletín CAÑAMIP 2004 Importancia del control de chinche salivosa (*Aeneolamia* spp) en gramíneas. No. 7. Julio.

CENGICANA (Centro Guatemalteco de Investigación y Capacitación de la Caña de azúcar). 2002. Informe anual 1998-1999. Guatemala. 58p.

Duque, M.C. 1996 Patrones de disposición espacial y su importancia en la definición de un plan de muestreo en MIP. Centro Internacional de Agricultura Tropical (CIAT) Cali, Colombia. 39 p.

Inventarios forestales para bosques latifoliados en América Central/ eds. Lorena Orozco, Cecilia Brumér. –Turrialba, C.R.: CATIE, 2002, 264 p.

Laurentin H, Pereira, C. 2002 Patrón de distribución y muestreo de estados inmaduros de Mosca Blanca, *Hemisia tabaci* (Gennadius) Homóptera: Aleyrodidae en Ajonjolí (*Sesamum indicum* L.). Bioagro 14(3): 145-152.

Malleux, J. 1982. Inventarios forestales en bosques tropicales. Universidad Agraria La Molina. Lima, Perú. 441 p.

Pantoja, A.; Fischer, A., Correa, F. 1997 Manejo Integrado de Plagas en Arroz. Centro Internacional de Agricultura Tropical (CIAT) Cali, Colombia. 114 p.

Scheaffer, R.L.; Mendenhall, W.; Ott, L. 1993 Elementos de muestreo. Trad. Rendón, G.; Gómez, J.R. Grupo Editorial Iberoamericana, México. 321 p.

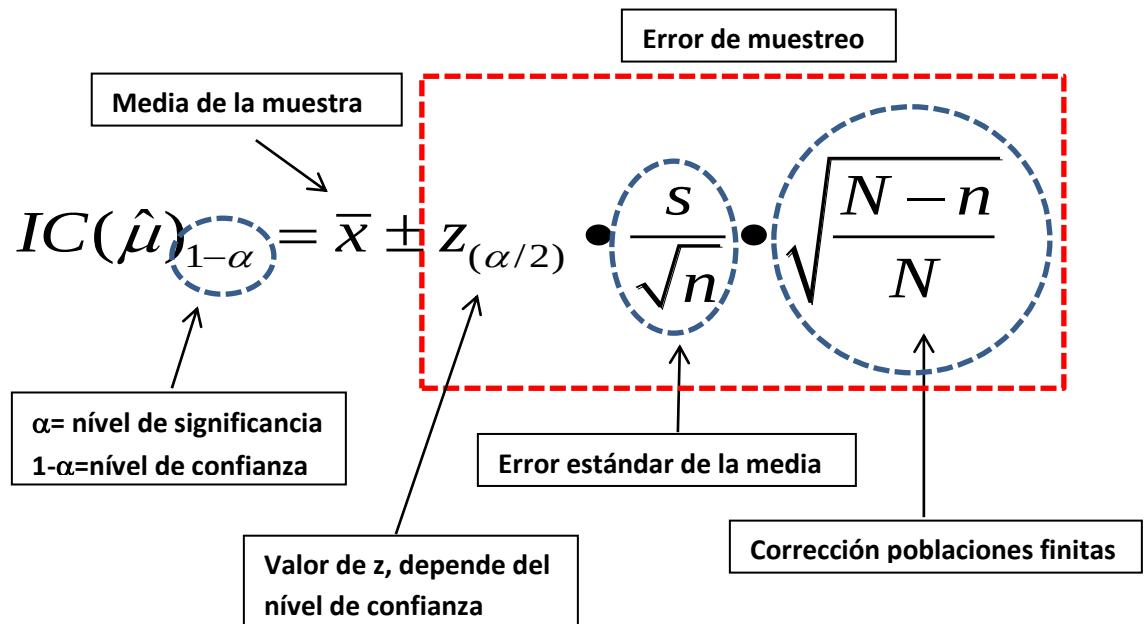
Stingel, E. 2005. Distribuição espacial e plano de amostragem para a cigarrinha-das-raízes, *Manaharva fimbriolata*, (Stal., 1854) em cana-de-açúcar. Disertación de Maestría. Universidad de São Paulo, Escuela Superior de Agricultura “Luiz de Queiroz”, Piracicaba. 74 p.

ANEXOS

DEDUCCION ECUACIONES UTILIZADAS EN MUESTREO SIMPLE ALEATORIO

Las ecuaciones para calcular el tamaño de muestra se deducen a partir de las ecuaciones de intervalo de confianza del parámetro que estamos interesados en estimar y si consideramos a la población como finita o infinita. A continuación se presenta la deducción de cada una de las ecuaciones para los parámetros: media (promedio), total y proporción.

ECUACIONES PARA ESTIMAR LA MEDIA



Caso 1: poblaciones infinitas

$$d = z_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}}$$

$$d^2 = z_{(\alpha/2)}^2 \cdot \frac{s^2}{n}$$

$$n = \frac{z_{(\alpha/2)}^2 \cdot s^2}{d^2}$$

Siendo d = error de muestreo, se obtiene multiplicando el valor de la media muestral por el valor de precisión (valores comunes oscilan entre 0.03-0.10)

Caso 2: poblaciones finitas

$$d = z_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}}$$

$$d^2 = z_{(\alpha/2)}^2 \cdot \frac{s^2}{n} \cdot \frac{N-n}{N}$$

$$N \cdot n \cdot d^2 = z_{(\alpha/2)}^2 \cdot s^2 \cdot (N-n)$$

$$N \cdot n \cdot d^2 = N \cdot z_{(\alpha/2)}^2 \cdot s^2 - n \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$N \cdot n \cdot d^2 + n \cdot z_{(\alpha/2)}^2 \cdot s^2 = N \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$n(N \cdot d^2 + z_{(\alpha/2)}^2 \cdot s^2) = N \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$n = \frac{N \cdot z_{(\alpha/2)}^2 \cdot s^2}{N \cdot d^2 + z_{(\alpha/2)}^2 \cdot s^2}$$

ECUACIÓN PARA ESTIMAR EL TOTAL

$$IC(\hat{\tau})_{1-\alpha} = N \left[\bar{x} \pm z_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} \right]$$

$$IC(\hat{\tau})_{1-\alpha} = N \cdot \bar{x} \pm N \cdot z_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}}$$

$$\hat{\tau} = N \cdot \bar{x}$$

Total estimado

Error de muestreo para el total

$$d = N \cdot \bar{x} \cdot \text{precisión}$$

Caso 1: poblaciones finitas

$$d = N \cdot z_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}}$$

$$d^2 = N^2 \cdot z_{(\alpha/2)}^2 \cdot \frac{s^2}{n} \cdot \frac{N-n}{N}$$

$$N \cdot n \cdot d^2 = N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2 \cdot (N-n)$$

$$N \cdot n \cdot d^2 = N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2 - n \cdot N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$N \cdot n \cdot d^2 + n \cdot N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2 = N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$n(N \cdot d^2 + N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2) = N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2$$

$$n = \frac{N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2}{N \cdot d^2 + N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2}$$

$$n = \frac{N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2}{N(d^2 + N \cdot z_{(\alpha/2)}^2 \cdot s^2)}, \text{ luego:}$$

$$n = \frac{N^3 \cdot z_{(\alpha/2)}^2 \cdot s^2}{N(d^2 + N \cdot z_{(\alpha/2)}^2 \cdot s^2)} \cdot \frac{\frac{1}{N}}{\frac{1}{N}}$$

Truco matemático

$$\Rightarrow n = \frac{N^2 \cdot z_{(\alpha/2)}^2 \cdot s^2}{d^2 + N \cdot z_{(\alpha/2)}^2 \cdot s^2}$$

ECUACIONES PARA ESTIMAR LA PROPORCIÓN

$$IC(\hat{p})_{1-\alpha} = p \pm z_{(\alpha/2)} \cdot \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N-n}{N}}$$

Caso 1: poblaciones infinitas

Error de muestreo
para la proporción

$$d = z_{(\alpha/2)} \cdot \sqrt{\frac{p \cdot q}{n}}$$

$$d^2 = z_{(\alpha/2)}^2 \cdot \frac{p \cdot q}{n}$$

$$n = \frac{z_{(\alpha/2)}^2 \cdot p \cdot q}{d^2}$$

Siendo **d** = error de muestreo, se ingresa directamente el valor de precisión (valores comunes oscilan entre 0.03-0.10)

Caso 2: poblaciones finitas

$$d = z_{(\alpha/2)} \cdot \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N-n}{N}}$$

$$d^2 = z_{(\alpha/2)}^2 \cdot \frac{p \cdot q}{n} \cdot \frac{N-n}{N}$$

$$N \cdot n \cdot d^2 = z_{(\alpha/2)}^2 \cdot p \cdot q \cdot (N-n)$$

$$N \cdot n \cdot d^2 = N \cdot z_{(\alpha/2)}^2 \cdot p \cdot q - n \cdot z_{(\alpha/2)}^2 \cdot p \cdot q$$

$$(N \cdot n \cdot d^2) + (n \cdot z_{(\alpha/2)}^2 \cdot p \cdot q) = N \cdot z_{(\alpha/2)}^2 \cdot p \cdot q$$

$$n(N \cdot d^2 + z_{(\alpha/2)}^2 \cdot p \cdot q) = N \cdot z_{(\alpha/2)}^2 \cdot p \cdot q$$

$$n = \frac{N \cdot z_{(\alpha/2)}^2 \cdot p \cdot q}{N \cdot d^2 + z_{(\alpha/2)}^2 \cdot p \cdot q}$$

ANEXO 2
ALGUNAS ECUACIONES PARA CALCULAR EL TAMAÑO DE MUESTRA EN EL CASO DE POBLACIONES INFINITAS

- a) Para estimado de intervalo de una media población.

$$n = \frac{z_{(\alpha/2)}^2 \times s^2}{E^2}, \text{ siendo } E = \text{error máximo permitido, siendo } E = \bar{x} \times d; d \text{ representa el valor de precisión: } 0.01, 0.05, 0.10, \text{ etc.}$$

Ejemplo: Si se desconoce el tamaño de la población (N) y se sabe que $s^2 = 462.8 \text{ u}^2$, el valor de la media es de: US\$ 422 y el valor de precisión es de 0.10, calcule el tamaño de la muestra, con un nivel de 95% de confianza.

$$n = \frac{1.96^2 \times 462.8}{(422 \times 0.10)^2} = 99.83 \approx 100 \text{ familias}$$

- b) Para estimado de intervalo de una proporción poblacional.

$$n = \frac{z_{(\alpha/2)}^2 \times p(1-p)}{E^2}, \text{ siendo } E = \text{error máximo permitido, por ejemplo: } 0.05, 0.10, 0.15, \text{ etc.}$$

Ejemplo: Si se desconoce el tamaño de la población (N) y se sabe que $p=0.5$ y el valor de precisión es de 0.05, calcule el tamaño de la muestra, con un nivel de 95% de confianza.

$$n = \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2} = 384.16 \approx 385 \text{ familias}$$